A CONCEPT TO ASSESS THE PERFORMANCE OF A PERMAFROST MODEL RUN FULLY COUPLED WITH A CLIMATE MODEL

A

DISSERTATION

Presented to the Faculty

of the University of Alaska Fairbanks

in Partial Fulfillment of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

By

Debasish PaiMazumder, B.Sc., M.Sc.

Fairbanks, Alaska

May 2009

A CONCEPT TO ASSESS THE PERFORMANCE OF A PERMAFROST MODEL RUN FULLY COUPLED WITH A CLIMATE MODEL

By

Debasish PaiMazumder

RECOMMENDED:

Gerhard Kramm

Uma S. Bhatt

John E. Walsh

Mike Masters

Advisory Committee Chair

Mike Masters

Chair, Department of Atmospheric Sciences

APPROVED:

Dean, College of Natural Science and Mathematics

Lawrence K. Duffy

Dean of the Graduate School

Apr 15, 2009

Date

Abstract

Soil-temperatures simulated by the fully coupled Community Climate System Model version 3.0 (CCSM3) are evaluated using three gridded Russian soil-temperature climatologies (1951-1980, 1961-1990, and 1971-2000) to assess the performance of permafrost and/or soil simulations. CCSM3 captures the annual phase of the soil-temperature cycle well, but not the amplitude. It provides slightly too high (low) soil-temperatures in winter (summer) with a better performance in summer than winter. In winter, soil-temperature biases reach up to 6 K.

Simulated near-surface air temperatures agree well with the near-surface air temperatures from reanalysis data. Discrepancies in CCSM3-simulated near-surface air temperatures significantly correlate with discrepancies in CCSM3-simulated soil-temperatures, i.e. contribute to discrepancy in soil-temperature simulation. Evaluation of cloud-fraction by means of the International Satellite Cloud Climatology project data reveals that errors in simulated cloud fraction explain some of the soil-temperature discrepancies in summer. Evaluation by means of the Global Precipitation Climatology Centre data identifies inaccurately-simulated precipitation as a contributor to underestimating summer soil-temperatures. Comparison to snow-depth observations shows that overestimating snow-depth leads to winter soil-temperature overestimation. Sensitivity studies reveal that uncertainty in mineral-soil composition notably contributes to discrepancies between CCSM3-simulated and observed soil-temperature climatology while differences between the assumed vegetation in CCSM3 and the actual vegetation in nature marginally contribute to the discrepancies in soil-temperature.

Out of the 6 K bias in CCSM3 soil-temperature simulation, about 2.5 K of the bias may result from the incorrect simulation of the observed forcing and about 2 K of the bias may be explained by uncertainties due network density in winter. This means that about 1.5 K winter-bias may result from measurement errors and/or model deficiencies.

Overall, the performance of a permafrost/soil model fully coupled with a climate model depends partly on the permafrost/soil model itself, the accuracy of the forcing data and design of observational network.

Table of Contents

Page

Page

Page

List of Figures

Page

List of Tables

Page

Acknowledgements

The breadth and depth of the study represents a collaboration and support from a variety of people. Due to their encouragement, support, teaching and financial support I was able to envision and explore subtle linkages between the different components of this study. At the very first opportunity I like to express my deep sense of gratitude to my advisor Nicole Mölders for her emancipated inspiration, continuous encouragement, sturdy and constructive ideas, revealing discussions and endless support. Her guidance and critique helped me conquer flaws in my style of thinking, writing and presentation. It really helped me to get more organized.

I take this opportunity to convey sincere thanks to my committee members US Bhatt, G Kramm and JE Walsh for their invaluable support and extension of all kinds of help whenever I needed.

Expressing gratitude in any form will be insufficient for the extensive co-operation and help I obtained from my research colleagues and junior and senior labmates. Special thanks are due to PA Bieniek, T Fathauer, SE Porter, R Tiruchirapalli, H Tran, TT Tran and MB Yarker for their moralizing encouragement and cooperation.

Computational support was provided in part by the National Center for Atmospheric Research and a grant of High Performance Computing (HPC) resources from the University of Alaska Fairbanks Arctic Region Supercomputing Center as part of the Department of Defense HPC Modernization Program. This thesis was funded by NSF under grant OPP-0327664 and ARC0652838.under the IARC cooperate agreement and EPSCoR-grant 0701898.

Finally, without the love, wishes and support from my family, this study would not have been possible. My heartfelt regards to my father Arin Kumar PaiMazumder, mother Mala Rani PaiMazumder, elder sister Tania PaiMazumder, uncle Himadri PaiMazumder and my wife Paramita Das. I dedicate my thesis to my grandfather Ajit Kumar PaiMazumder for his constant support and encouragement in every step of my life. The spirit of the work also came from the bewildering natural beauty of Alaska. For me, crossing the latitude proved very rewarding!

Chapter 1 Introduction

The better understanding of feedbacks between permafrost, weather, climate, and other potential impacts like economic and infrastructure damages, ecosystem changes, and freshwater availability (Esch and Osterkamp 1990, Cherkauer and Lettenmaier 1999, Oechel et al. 2000, Serreze et al. 2000, Zhuang et al. 2001) requires adequate soil-temperature simulations in numerical weather prediction models (NWPMs), chemistry transport models (CTMs), General Circulation Models (GCMs), and Earth System Models (ESMs). Soil-temperature is typically simulated by a soil or permafrost model embedded in NWPMs, CTMs, GCMs and ESMs (Robock et al. 1995, Kramm et al. 1996, Huang et al. 1996, Mölders and Romanovsky 2006, Narapusetty and Mölders 2006, Mölders and Kramm 2009). Permafrost or soil models are run at each land point as a column; thus they are typically evaluated in an offline mode driven by observations (Mölders and Romanovsky 2006, Narapusetty and Mölders 2006). The coupling of a more realistic permafrost, soil or Land Surface Models (LSMs) to a climate model offers the prospect of improved simulation skill. In this context, the great achievements of the Project for Intercomparison of Land Surface Parameterization Schemes (PILPS) need to be discussed (e.g. Henderson-Sellers et al. 1993, 1995, Shao and Henderson-Sellers 1996, Qu et al. 1998, Schlosser et al. 2000, Slater et al. 2001, Luo et al. 2003).

In order to improve the parameterization of land-surface processes, especially hydrological, energy, momentum and carbon exchanges between the atmosphere and the continental surface, PILPS was initiated in 1992 as a World Climate Research Programme project. The goal was to enhance understanding of the differences in LSM performance as they are used in climate and weather prediction models (Henderson-Sellers et al. 1993, 1995, Qu et al. 1998). In PILPS' first phase, offline simulations of various LSMs were compared; the main differences found were spin-up time and flux partitioning (Henderson-Sellers et al. 1993, 1995). A main result of PILPS' first phase was that spin-up time for various land-surface schemes depends on the scheme itself, total moisture holding capacity and initialization of moisture storage (Yang and

Dickinson 1995). PILPS' second phase evaluated the offline performances of various LSMs against observations (Schlosser et al. 2000, Slater et al. 2001, Luo et al. 2003). A main finding was that average annual water and energy partitioning differed appreciably among schemes. PILPS studies focusing on LSM performance at high latitudes revealed that peat-lands and bogs introduce errors (Bowling et al. 2003). Due to LSM complexities, it is difficult to generalize causes for any differences (Nijssen et al. 2001). However, inclusion of frozen-ground physics was identified as an important aspect to simulate energy and water fluxes at the earth's surface (e.g. Cherkauer and Lettenmaier 1999) as well as soil-temperature and soil-moisture states (Montaldo and Albertson 2001). Capturing the freezing front depth is also important because temperature variations diminish greatly in the deeper soil (Luo et al. 2003, Mölders and Romanovsky 2006).

PILPS' third phase involved the evaluation of LSM simulations in coupled mode with a global atmospheric circulation models jointly with the Atmospheric Model Intercomparison Project (AMIP) because errors in simulated atmospheric forcing can affect near-surface conditions with further feedback to simulated soil conditions (Henderson-Sellers et al. 1995, Chen and Dudhia 2001, Zeng et al. 2002, Narapusetty and Mölders 2005). The major finding was that LSMs captured soil-temperature conditions better when run in an offline mode with known site-specific parameters and observed atmospheric forcing rather than in a coupled mode; it must be re-evaluated when run in fully-coupled mode, which lead to the setup of PILPS' fourth phase (Qu and Henderson-Sellers 1998). PILPS' fourth phase evaluated land-surface parameterization schemes fully coupled to atmospheric models. Narapusetty and Mölders (2005) showed that model results may vary depending on the parameterizations chosen to simulate the atmospheric forcing.

However, if LSM is implemented into a NWPM, CTM, ESM, or GCM, it must be re-evaluated because fully-coupled models introduce additional uncertainty and error sources. The large grid-cells of GCMs and ESMs do not permit direct evaluation because several sites may exist within one grid-cell. Choice of initial conditions, discretization,

grid resolution, numerical scheme, parameterizations, model assumptions, and/or empirical parameters may cause an incorrect simulation of atmospheric forcing (Anthes et al. 1989, Slater et al. 1998, Mölders 2005, Mölders et al. 2005, Narapusetty and Mölders 2005, 2006). Other sources of errors could be in the initialization of soil-moisture and temperature distribution, parameterization of sub-grid scale processes, surface run-off and cloud microphysical processes, which have been investigated by various authors (Avissar and Pielke 1989, Calder et al. 1995, Mölders et al. 1996, 1997, 2003, Wang and Kumar 1998, Douville and Chauvin 2000, Niu and Yang 2004, Narapusetty and Mölders 2005). Uncertainties in land-use, soil types and vegetation can lead to errors in near-surface air temperature, humidity and soil fluxes (Mölders et al. 1997, Mölders 2001, 2005, Mölders et al. 2005).

GCMs provide volume averages of simulated variables with area averages of several 100 square-kilometers in the horizontal and several decameters in the vertical direction. Thus, they are difficult to compare to measurements at a site (point measurement). Therefore, using interpolation of available measurements to the climate model grid has become common practice (Palutikof et al. 1997, Bauer et al. 2002, Li et al. 2008). It is expected that such interpolation may introduce uncertainty into the grid-cell averages and, hence, any evaluation. Uncertainty in gridded regional averages has been investigated with respect to the interpolation methods for precipitation, radiation, air pollutants and meteorological state variables (Shaw and Lynn 1972, Creutin and Obled 1982, Court and Bare 1984, Lebel et al. 1987, Lindley and Walsh 2004, Luo et al. 2008).

Observation network density and design may introduce some uncertainty in the regional averages. Pielke et al. (2007) found that near-surface temperatures derived from poorly and non-homogeneously sited stations differ more greatly than well-sited stations. Changes in site location or network density alter the topography, latitude and elevation represented by the network; these changes influence air-temperature and minimum and maximum temperatures measured at these sites (Robeson and Doty 2005, Peterson 2006). Network density may also affect regional precipitation averages (Frei and Schär 1998, Tsintikidis et al. 2002); high-density networks are more likely to capture locally high

precipitation rates than coarse networks (St.-Hilarie et al. 2003). Urbanization, land-cover changes, moving, shutting down or adding of sites, errors in digitizing old paper records, the procedure of filling missing data and, to a certain degree, the applied interpolation algorithms may influence the accuracy and reliability of long time-series of gridded data compiled from all available stations (Mitchell et al. 2004).

Other sources of errors include the parameterization of the permafrost model itself and the discretization and numerical scheme. For example, Mölders and Romanovsky (2006) found that the choice of soil thermal conductivity parameterization, lower boundary condition and vertical grid resolution have greatest influence on the accuracy of simulated soil-temperature. Narapusetty and Mölders (2006) revealed that a Galerkin weak finite element numerical scheme is more suitable in comparison to Crank-Nicholson finite-difference scheme in the soil model of the Hydro-Thermodynamics Soil-Vegetation Scheme (Kramm et al. 1996, Mölders et al. 2003). Nicolsky et al. (2007) also found that modification to the numerical scheme in a permafrost model improve the simulation of permafrost dynamics. Different configurations of soil layers in permafrost or soil model also can introduce biases in permafrost simulation (Mölders and Romanovsky 2006, Alexeev et al. 2007).

The present study assesses the performance of a permafrost model (Dai et al. 2003, Oleson et al. 2004, Dickinson et al. 2006) fully coupled with a climate model. The research hypothesis of this dissertation is that *the performance of a permafrost model fully coupled with a climate model will depend on the permafrost model itself as well as on the accuracy of the forcing data and the data for evaluation of soil-temperature that may limit full assessment of the performance.* To test this hypothesis, three specific questions will be answered:

a. How well does a fully coupled climate model simulate soil-temperature?
b. What are the reasons for the discrepancies between simulated and observed soil-temperature?

c. How much of the discrepancies between simulated and observation-based gridded soil-temperature climatologies can be explained by the observational network density and/or design?

In this study, simulations are performed with the fully coupled Community Climate System Model version 3 (CCSM3; Collins et al. 2006) for 50 years and analyzed by means of three gridded climatologies of soil-temperature data (1951-1980, 1961-1990, 1971-2000) based on observational data. Furthermore, two CCSM3 simulations wherein the sand fraction is reduced (enhanced) by 10 % and the clay fraction is enhanced (reduced) by 10 % in Russia are performed and analyzed to assess the impact of soil-parameter inaccuracy on CCSM3-simulated soil-temperatures and to help in answering questions (a and b) above.

To investigate the impact of network density and design on regional averages (question c), simulations are performed with the Weather Research and Forecasting (WRF; Skamarock et al. 2005) model over Russia for July and December 2005, 2006 and 2007 to create a "dataset". Based on the values at all WRF grid-points, regional averages for various quantities are calculated for $2.8^{\circ} \times 2.8^{\circ}$ areas as the "reference". Regional averages determined based on 40 artificial networks with ten set of 500, 400, 200, or 100 different randomly chosen WRF-grid-points as "sites" as well as 411 "sites" that correspond to the locations of a real or historic network, are compared with the reference regional averages.

Chapter 2 primarily discusses the physical packages of CCSM3 relevant for evaluation of soil-temperature by gridded soil-temperature and assessing the impact of simulated atmospheric forcing to determine overall error, the focal aspect of this dissertation. In chapter 3, the methodology of this research, including descriptions of the various CCSM3 simulations, is described in detail. In chapter 4, results of the performance of CCSM3 in simulating soil-temperature and the contribution of inaccurate simulation of various quantities to the discrepancies between CCSM3-simulated and observed soil-temperature are presented to address questions a and b concerning CCSM3's performance. In chapter 5, the assessment of the uncertainty in regional

averages of various quantities is discussed in detail. Chapter 6 serves to answer question c, where accuracy of gridded soil-temperature data based on available station data and development of the recommendation for network design to optimize their use for model validation are explained. The overall conclusions are presented in chapter 7.

References

Alexeev VA, Nicolsky DJ, Romanovsky VE, Lawrence DM (2007) An evaluation of deep soil configuration in the CLM3 for improve representation of permafrost. Geophys Res Lett 34: L09502, doi:10.1029/2007GL029536

Anthes RA, Kuo YH, Hsie EY, Low-Nam S, Bettge TW (1989) Estimation of skill and uncertainty in regional numerical models. Quart J Roy Meteorol Soc 111: 763-806

Avissar R, Pielke RA (1989) A parameterization of heterogeneous land surface for atmospheric numerical models and its impact on regional meteorology. Mon Wea Rev 117: 2113-2136

Bauer M., Del Genio AD, Lanzante JR (2002) Observed and simulated temperature-humidity relationship: Sensitivity to sampling analysis. J Climate 15: 203-215

Bowling LC, Lettenmaier DP, Nijssen B, Graham LP, Clark DB, Maayar M, Essery R, Goers S, Gusev YM, Habets F, Hurk B, Jin J, Kahan D, Lohmann D, Ma X, Mahanama S, Mocko D, Nasonova O, Niu GY, Samuelsson P, Shmakin AB, Takata K, Verseghy D, Viterbo P, Xia Y, Xue Y, Yang ZL (2003) Simulation of high latitude hydrological processes in the Torne-Kalix basin: PILPS Phase 2(e) 1: Experiment description and summary intercomparisons. Global and Planetary Change 38: 1-30

Calder IR, Hall RA, Bastable HG, Gunston HM, Shela O, Chirwa A, Kafundu R (1995) The impact of land use changes on water resources in sub-Saharan Africa: A modelling study of Lake Malawi. J Hydrol 170: 123-136

Chen F, Dudhia J (2001) Coupling an advanced land surface-hydrology model with the Penn State-NCAR Modeling System. Part I: Model implementation and sensitivity. Mon Wea Rev 129: 569-585

Cherkauer KA, Lettenmaier DP (1999) Hydrologic effects of frozen soils in the upper Mississippi River basin. J Geophys Res 104D: 19599-19610

Collins WD, Bitz CM, Blackmon ML, Bonan GB, Bretherton CS, Carton JA, Chang P, Doney SC, Hack JJ, Henderson TB, Kiehl JT, Large WG, McKenna DS, Santer BD, Smith RD (2006) The Community Climate System Model: CCSM3. J Climate 19: 2122-2143

Court A., Bare MT (1984) Basin precipitation estimates by Bethlahmy's two-axis method. J Hydrol 68: 149-158

Creutin JD, Obled C (1982) Objective analysis and mapping techniques for rainfall fields: an objective comparison. Water Resour Res 15: 1752-1762

Dai Y, Zeng X, Dickenson RE, Backer I, Bonan GB, Bosilovich MG, Denning AS, Dirmeyer PA, Houser PR, Niu G, Oleson KW, Schlosser CA, Yang ZL (2003) The Common Land Model. Bull Amer Meteor Soc 84: 1013-1023

Dickinson RE, Oleson KW, Bonan G, Hoffman F, Thornton P, Vertenstein M, Yang Z-L, Zeng X (2006) The Community Land Model and its climate statistics as a component of the Community Climate System Model. J Climate 19: 2302-2324

Douville H, Chauvin F (2000) Relevance of soil moisture for seasonal climate predictions: A preliminary study. Climate Dyn 16: 719-736

Esch DC, Osterkamp TE (1990) Cold regions engineering: climatic warming concerns for Alaska. J Cold Regions Engineering 4: 6-14

Frei C, Schär C (1998) A precipitation climatology of the Alps from high resolution rain-gauge observation. Int J Climatol 18: 873-900

Henderson-Sellers A, Yang ZL, Dickinson RE (1993) The Project for Intercomparison of Land-surface Parameterization Schemes. Bull. Am. Meteorol. Soc. 74: 1335-1349

Henderson-Sellers A, Pitman AJ, Love PK, Irannedjad P, Chen TH (1995) The Project for Intercomparison of Land surface Parameterization Schemes (PILPS): Phase 2 and 3. Bull Amer Meteorol Soc 76: 489-503

Huang J, Van den Dool HM, Georgakakos KP (1996) Analysis of model-calculated soil moisture over United States (1931-1993) and applications to long-range temperature forecast. J Climate 9: 1350-1362

Kramm G, Beier N, Foken T, Müller H, Schröder P, Seiler W (1996) A SVAT scheme for NO, NO2, and O3- Model description and test results. Meteorol Atmos Phys 61: 89-106

Lebel T, Bastin G, Obled C, Creutin JD (1987) On the accuracy of areal rainfall estimation: a case study. Water Resour Res 23: 2123-2134

Li, Z, Bhatt US, Mölders N (2008) Impact of doubled CO_2 on the interaction between the regional and global water cycle in four study regions. Clim Dyn 30: 255-275

Lindley SJ, Walsh T (2004) Inter-comparison of interpolated background nitrogen dioxide concentrations across Greater Manchester, UK. Atmos Environ 39: 2709-2724

Luo L, Robock A, Vinnikov KY, Schlosser CA, Slater AG, Boone A, Braden H, Cox P, deRosany P, Dickinson E, Dai Y, Diuan Q, Etchevers P, Henderson-Sellers A, Gedney N, Gusev YM, Habets F, Kim J, Kowalczyk E, Mitchell K, Nasonova ON, Noilhan J, Pitman AJ, Schaake J, Shmakin AB, Smirnova TG, Wetzel P, Xue Y, Yang Z-L, Zeng QC (2003) Effects of frozen soil on soil temperature, spring infiltration, and runoff: Results from the PILPS2(d) experiment at Valdai, Russia. J Hydrometeor 4: 334-351

Luo W, Taylor MC, Parker SR (2008) A comparison of spatial interpolation methods to estimate continuous wind-speed surfaces using irregularly distributed data from England and Wales. Int J Climatol 28: 947-959

Mitchell TD, Carter TR, Jones PD, Hulme M, New M (2004) A comprehensive set of high-resolution grids of the monthly climate for Europe and the globe: the observed record (1901-2000) and 16 scenarios (2002-2100). Working paper 55, Tyndall Centre for Climate Change Research, Norwich, pp

Mölders N, Raabe A, Tetzlaff G (1996) A comparison of two strategies on land surface heterogeneity used in a mesoscale beta meteorological model. Tellus 48A: 733-749

Mölders N, Strasser U, Schneider K, Mauser W, Raabe A (1997) A sensitivity study on the initialization of surface characteristics in meso-β/γ-modeling using digitized vs. satellite derived landuse data Contrib Atmos Phys 70 173-187

Mölders N (2001) On the uncertainty in mesoscale modeling caused by surface parameters. Meteor Atmos Phy 76: 119-141

Mölders N, Haferkorn U, Döring J, Kramm G (2003) Long-term numerical investigations on the water budget quantities predicted by the hydro-thermodynamic soil vegetation scheme (HTSVS) - Part II: Evaluation, sensitivity, and uncertainty. Meteor Atmos Phys 84: 137-156

Mölders N (2005) Plant- and soil parameter-caused uncertainty of predicted surface fluxes. Mon Wea Rev 133: 3498-3516

Mölders N, Jankov M, Kramm G (2005) Application of Gaussian error propagation principles for theoretical assessment of model uncertainty in simulated soil processes caused by thermal and hydraulic parameters. J. Hydrometeorol 6: 1045-1062

Mölders N, Romanovsky VE (2006) Long-term evaluation of the Hydro-Thermodynamic Soil-Vegetation Scheme's frozen ground/permafrost component using observations at Barrow, Alaska. J Geophys Res 111: D04105, doi: 10.1029/2005JD005957

Mölders N, Kramm G (2009) Permafrost modeling in weather forecasts and climate projections. In: Permafrost: Impact, Properties and Engineering, Nova Science

Montaldo N, and Albertson JD (2001) On the use of the force-restore SVAT model formulation for stratified soils. J. Hydrometeor. 2: 571-578

Narapusetty B, Mölders N (2005) Evaluation of snow depth and soil temperatures predicted by the Hydro-Thermodynamic Soil-Vegetation Scheme (HTSVS) coupled with the PennState/NCAR Mesoscale Meteorological Model (MM5). J Appl Meteorol 44: 1827-1843

Narapusetty B, Mölders N (2006) Evaluation of the soil module of HTSVS by observations and a theoretically advanced numerical scheme. Mon Wea Rev 134: 2927-2942

Nicolsky DJ, Romanovsky VE, Alexeev VA, Lawrence DM (2007) Improved modeling of permafrost dynamics in a GCM land-surface scheme. Geophys Res Lett 34: L08501, doi:10.1029/2007GL029525

Nijssen B, Schnur R, Lettenmaier DP (2001) Global retrospective estimation of soil moisture using the variable infiltration capacity Land Surface Model, 1980-93. J. Climate 14: 1790-1808

Niu GY, Yang ZL (2004) The effects of canopy processes on the snow surface energy and mass balances. J Geophys Res 109, D23111, doi: 10.1029/2004JD004884

Oechel WC, Vourlitis GL, Hastings SJ, Zulueta RC, Hinzman L, Kane D (2000) Acclimation of ecosystem CO_2 exchange in the Alaskan Arctic in response to decadal climate warming. Nature 406: 978-981

Oleson KW, Dai Y, Bonan GB, Bosilovich M, Dickinson R, Dirmeyer P, Hoffman F, Houser P, Levis S, Niu GY, Thornton P, Vertenstein M, Yang ZL, Zeng X (2004) Technical description of Community Land Model (CLM). Technical Report NCAR/TN-461+STR, National Center for Atmospheric Research, Boulder, CO. 80307-3000, pp 174

Palutikof JP, Winkler JA, Goodess CM, Andresen JA (1997) The simulation of daily temperature time series from GCM output. Part I: Comparison of model data with observation. J Climate 10: 2497-2531

Peterson TC (2006) Examination of potential biases in air temperature caused by poor station locations. Bull Amer Meteorol Soc 87: 1073-1080

Pielke Sr RA, Nielsen-Gammon J, Davey C, Angel J, Bliss O, Cai M, Doesken N, Fall S, Niyogi D, Gallo K, Hale R, Hubbard KG, Lin X, Li H, Raman S (2007) Documentation of uncertainties and biases associated with surface temperature measurement sites for climate change assessment. Bull Amer Meteorol Soc 88: 913-928

Qu W, Henderson-Sellers A (1998) Comparing the scatter in PILPS off-line experiments with those in AMIP I coupled experiments. Global Planetary Change 19: 209-224

Qu W, Henderson-Sellers A, Pitman AJ, Chen TH, Abramopoulos F, Boone A, Chang S, Chen F, Dai Y, Dickinson RE, Dumenil L, Ek M, Gedney N, Gusev YM, Kim J, Koster R, Kowalczyk EA, Lean J, Lettenmaier D, Liang X, Mahfouf JF, Mengelkamp HT, Mitchell K, Nasonova ON, Noilhan J, Robock A, Rosenzweig C, Schaake J, Schlosser CA, Schulz JP, Shmakin AB, Verseghy DL, Wetzel P, Wood EF, Yang ZL, Zeng Q (1998) Sensitivity of latent heat flux from PILPS land-surface schemes to perturbations of surface air temperature. J Atmos Sci 55: 1909-1927

Robeson SM, Doty JA (2005) Identifying rogue air temperature station using cluster analysis of percentile trends. J Climate 18: 1275-1287

Robock A, Konstantin YV, Schlosser CA, Nina AS, Yongkang X (1995) Use of midlatitude soil moisture and meteorological observations to validate soil moisture simulations with biosphere and bucket models. J Climate 8: 15-35

Schlosser CA, Slater AG, Robock A, Pitman AJ, Vinnikov KY, Henderson-Sellers A , Speranaskaya NA, Mitchell K and the PILPS 2(D) contributors (2000): Simulation of a Boreal grass land hydrology at Valdai, Russia: PILPS phase 2(d). MonWea Rev 128: 301-321

Serreze MC, Walsh JE, Chapin FS, Osterkamp TE, Dyurgerov M, Romanovsky V, Oechel WC, Morison J, Zhang T, Barry RG (2000) Observational Evidence of Recent Change in the Northern High-Latitude Environment. Climatic Change 46: 159-207

Shao Y, Henderson-Sellers A (1996) Modeling soil moisture: A project for intercomparison of land surface parameterization schemes phase 2 (b). J Geophys Res 101D: 7227-7250

Shaw EM, Lynn PP (1972) A real rainfall evaluation using two surface fitting techniques. Hydrol Sci Bull 17: 419-433

Skamarock WC, Klemp JB, Dudhia J, Gill DO, Baker DM, Wang W, Powers JG (2005) A description of the advanced research WRF version 2. NCAR Technical Note, NCAR/TN-468+STR, pp. 88

Slater AG, Pitman AJ, Desborough CE (1998) The validation of snow parameterization designed for use in general circulation models. Int J Climatol 18: 595-617

Slater AG, Schlosser CA, Desborough CE, Pitman AJ, Henderson-Sellers A, Robock A, Vinnikov KY, Mitchell K, Boone A, Braden H, Chen F, Cox P, deRosany P, Dickinson E, Dai Y, Diuan Q, Entin J, Etchevers P, Gedney N, Gusev YM, Habets F, Kim J, Koren V, Kowalczyk E, Nasonova ON, Noilhan J, Schaake J, Shmakin AB, Smirnova TG, Verseghy D, Wetzel P, Xue Y, Yang Z-L, Zeng QC (2001) The Representation of Snow in Land Surface Scheme: Results from PILPS 2(d). Bull Amer Meteorol Soc 2: 7-25

St.-Hilaire A, Taha BM, Ouarda J, Lachance M, Bobée B, Gaudet J, Gignac C (2003) Assessment of the impact of meteorological network density on the estimation of basin precipitation and runoff: a case study. Hydrol Process, 17: 3561-3580

Tsintikidis D, Georgakakos KP, Sperfslage JA, Smith DE, Carpenter TM (2002) Precipitation uncertainty and rain gauge network design within Folsom Lake watershed. J Hydrologic Engrg 7: 175-184

Wang W, Kumar A (1998) A GCM assessment of atmospheric seasonal predictability associated with soil moisture anomalies over North America. J Geophys Res 103: 28637-28646

Yang ZL, Dickinson RE (1995) Preliminary study of spin-up processes in land surface models with the first stage data of Project for Intercomparison of Land Surface Parameterization Schemes Phase 1 (a). J Geophys Res 100: 16553-16578

Zeng X, Shaikh M, Dai Y, Dickinson RE, Myneni R (2002) Coupling of the Common Land Model to the NCAR Community Climate Model. J Climate 15: 1832-1854

Zhuang Q, Romanovsky VE, McGuire AD (2001) Incorporation of a permafrost model into a large-scale ecosystem model: Evaluation of temporal and spatial scaling issues in simulating soil thermal dynamics. J Geophys Res 106D: 33,649-33,670

Chapter 2 Model description

2.1 Description of Community Climate System Model version 3 (CCSM3)

2.1.1 Introduction

The third version of the coupled global climate system model, Community Climate System Model (CCSM3; Collins et al. 2006a) is used in this study. The CCSM3 consists of the Community Atmosphere Model version 3 (CAM3; Collins et al. 2006b), the Common Land Model version 3 (CLM3; Dai et al. 2003, Oleson et al. 2004, Dickinson et al. 2006), the Community Sea Ice Model version 5 (CSIM5; Briegleb et al. 2004) and the Parallel Ocean Program version 1.4.3 (POP1.4.3; Smith et al. 1992). In the CCSM3 simulations, these four components exchange information through a coupler (cpl) without flux correction. The CCSM3 better simulates the surface temperature in winter, tropopause temperature and different atmospheric forcing in tropics in comparison to previous versions of the model (Collins et al. 2006a).

The following sections primarily discuss those physical packages of CCSM3 that are relevant for the assessment of the performance of a permafrost model fully coupled with a climate model including those parameterizations relevant for the atmospheric forcing, focal aspect of this dissertation. For further details on CCSM3 see Smith et al. (1992), Oleson et al. 2004, Dai et al. (2003), Collins et al. (2004) and Briegleb et al. (2004).

2.1.2 Community Atmospheric Model version 3

CAM3 (Collins et al. 2004) is the latest version of NCAR atmospheric general circulation models (AGCMs). CAM3 can be run in standalone mode and as a component of CCSM3 (Collins et al. 2004). CAM3 is designed to support T85, T42, and T31 configurations. The configuration used here employs a T42 configuration (2.8^{o}-resolution) with 64x128 grid cells in the zonal and meridional directions, respectively, and 26 layers in the vertical direction is employed in the present study. It employs a hybrid terrain-following sigma-coordinate system.

2.1.2.1 Dynamics

CAM3 consists of two parts: the dynamic core and parameterization suite. Both parts are coupled to each other using time and operator splitting (Collins et al. 2004). The dynamical core can be determined by Eulerian, semi-Lagrangian, or finite-volume methods. In this study, the Eulerian method is applied.

CAM3 uses the prognostic equations for the momentum, temperature and moisture and the continuity equation. The momentum equation is represented in terms of divergence, vorticity and the horizontal diffusion. The temperature equation is expressed by divergence, vertical advection, sources and sinks and the horizontal diffusion of temperature. The moisture equation includes divergence, vertical advection and sources and sinks of moisture (Collins et al. 2004). The source and sink terms in the prognostic equations are estimated from diagnostic equations. All the prognostic equations with their boundary conditions are solved in time using a semi-implicit leapfrog scheme. In this scheme, the linear, nonlinear and vertical terms of the prognostic equations can be approximated by the averaging between two consecutive time steps, central difference and finite difference methods, respectively. The horizontal diffusion terms are excluded in this solution method of solving the equations. CAM3 includes the linear horizontal diffusion terms for prognostic variables to produce the kinetic energy spectra. The horizontal diffusion operator in CAM3 is applied to ensure that the Courant-Friedrichs-Levy (CFL) criterion is fulfilled in the upper layer of the model (Collins et al. 2004). The horizontal diffusion terms are calculated implicitly in the spectral space.

2.1.2.2 Physics

CAM3 considers cloud and precipitation processes, long-wave and shortwave radiation processes, turbulence and exchange of heat and moisture at the atmosphere-vegetation interface. All of the aforementioned processes are parameterized. The following sections discuss those parameterization schemes of CAM3 that are important for this study.

2.1.2.2.1 Precipitation processes

The various aspects of cloud and precipitation formation processes are considered by the deep convection scheme (Zhang and McFarlane 1995), shallow convection scheme (Hack 1994), evaporation of convective precipitation (Sundqvist 1988), bulk-microphysics parameterization (Rasch and Kristjánsson 1998, Zhang et al. 2003) and cloud-fraction parameterization (Slingo 1987, Hack et al. 1993).

2.1.2.2.1.1 Deep convection

In the deep convection scheme, an air parcel moves vertically from the lower troposphere to about 500hPa. In CAM3, the deep convective scheme is based on the plume ensemble approach (Zhang and McFarlane 1995, Collins et al. 2004). The time rate of change in temperature and moisture of the cloud layer are represented by vertical cloud-mass flux including updraft and downdraft, the dry static energy and specific humidity.

The updraft is expressed as a collection of entraining plumes, each with a characteristic entrainment rate. At the top of the plume, the updrafts detrain the mass into the environment in a thin layer. The entrainment rate for the plume is estimated by solving the moist static energy equation iteratively at each level with known lower boundary condition. The cloud-updraft mass flux and the updraft-budget equation for water-vapor mixing ratio, moist and dry static energy and cloud liquid water can be expressed by the cloud-base mass flux, entrainment rate, detrainment rate and cloud-base height by assuming the same cloud-base mass flux for all plumes in the vertical direction.

Downdrafts are initiated at or below the bottom of the updraft detrainment layer due to the production of precipitation. Downdraft mass flux can be expressed similarly to that of the updraft, but it contains a proportionality factor (function of total precipitation in the convective layer and rainwater evaporation needed to retain the downdraft) to ensure the physical consistency between the downdraft strength and precipitation availability.

The deep convective scheme considers updrafts and downdrafts and relates cloud-base mass flux to the Convective Available Potential Energy (CAPE) (Collins et al.

2004). The change in CAPE can be expressed as a function of CAPE consumption rate per unit cloud-base mass flux due to the linear relationship between cloud-base mass flux and large-scale temperature and moisture changes in cloud layer. In CAM3, cloud water and hydrometeors are transported by deep convection (Zhang and McFarlane 1995, Collins et al. 2004).

2.1.2.2.1.2 Shallow convection

In the shallow convection scheme, it is assumed that vertical lifting is limited to below 500hPa; convection is described by a vertically discrete non-entraining cloud model with three levels. The non-entraining convective element enters in the lower level and due to unstable conditions it rises to the mid-troposphere where condensation and rainwater formation occur. Detrainment occurs in the top level (Collins et al. 2004). The large-scale budget equations for dry static energy and the specific humidity for these three layers can be represented as a function of convective mass flux at cloud base, the convective scale of the liquid water sink and detrainment parameter at the center of the top and mid-troposheric levels (Yanai et al. 1973, Collins et al. 2004). The detrainment parameter varies from 0 to 1. It is obtained from a buoyancy argument that estimates the actual auto-conversion from cloud-water to rainwater. The detrainment parameter reaches its maximum value for minimum auto-conversion requirement. The convective mass flux is determined by detrainment parameter for each level. Consequently, a vertical profile of the total cloud-mass flux associated with the shallow convection can be constructed.

The moist and dry static energy and specific humidity of the updraft region are equal to the sum of their large-scale values in the sub-cloud layer within the diagnosed atmospheric boundary layer (ABL) and arbitrary thermodynamic perturbations. The perturbation quantities for specific humidity and static energy are equal to temperature excess (strength of convective thermals in the lower part of the ABL) and zero, respectively (Collins et al. 2004).

2.1.2.2.1.3 Evaporation of convective precipitation

The evaporation rate of convective precipitation is a function of local large-scale sub-saturation and the rate of convective rainwater flux (Sundqvist 1988). It is used to evaluate the local change in specific humidity and temperature at each layer.

2.1.2.2.1.4 Cloud microphysical parameterization

The parameterization of resolvable scale cloud microphysical processes follows Rasch and Kristjánsson (1998) and Zhang et al. (2003). It estimates the condensation with the change in cloud fraction and in-cloud condensate (Collins et al. 2004). The parameterization consists of two main components:

- The macro-scale component for exchanging water substance between the condensate and vapor phase (Zhang et al. 2003), and
- the bulk-microphysical component for considering conversion from condensate to precipitating particles (Rasch and Kristjánsson 1998)

In the macro-scale component, convection influences cloud-cover through convective tendencies that include evaporation of the convective cloud and convective precipitation. Water vapor, temperature and cloud water tendencies are estimated from the detrainment of evaporating cloud-water in accordance with Zhang and McFarlane (1995).

In the bulk-microphysical component, precipitation formation occurs due to (1) conversion of cloud-water to rainwater (Chen and Cotton 1987), (2) collection of cloud-water by rainwater from above (Tripoli and Cotton 1980), (3) auto-conversion of ice to snow (Kessler 1969, Lin et al. 1983, Sundqvist 1988), (4) collection of ice by snow (Lin et al. 1983) and (5) collection of liquid particles by snow (Lin et al. 1983).

2.1.2.2.1.5 Cloud-fraction parameterization

Cloud fraction and cloud optical properties of low level marine stratus, convective cloud and layered cloud are diagnosed in accord with Slingo (1987), Hack et al. (1993),

Kiehl et al. (1998) and Rasch and Kristjánsson (1998). Total cloud cover is represented by a maximum overlap assumption of cloud types within each grid-cell.

2.1.2.2.2 Radiation parameterization

The long-wave and shortwave heating rates are calculated hourly with constant long-wave and shortwave flux and flux divergence between hours. The insolation at the top of the atmosphere is defined as a function of the solar constant, solar zenith angle and the distance between the earth and the sun (Berger 1978, Collins et al. 2004).

The shortwave radiative parameterization includes cloud vertical overlap (Collins 2001), near-infrared absorption by water vapor, and prescribed aerosol datasets for consideration of shortwave aerosol radiative forcing. The atmosphere is comprised of a vertical set of horizontal homogeneous layers with a specified heating rate (Collins et al. 2004). Surface albedo distinguishes albedo for direct and diffuse incident radiation and albedo differs depending on the surface (Collins et al. 2004). Clear-sky and overcast sky flux and heating are also determined (Kiehl and Ramanathan 1990, Collins et al. 2004).

CAM3 considers two types of aerosols: soil dust and carbonaceous. CAM3 considers direct and semi-direct effect of aerosols on shortwave fluxes and heating rates, but it neglects the indirect effect discussed by Twomey et al. (1984). The aerosol climatological dataset stems from the Model for Atmospheric Chemistry and Transport (MATCH; Rasch et al. 1997) and satellite-estimated aerosol optical depth from NOAA Pathfinder II dataset (Stowe et al. 1997).

Cloud drop effective radius is assumed to be 14μm over ocean and sea-ice. Over land, cloud drop radius varies from 8μm to 14μm and also depends on temperature (Collins et al. 2004).

The broad band model approach introduced by Ramanathan (1976) and modified by Kiehl and Briegleb (1991) and Kiehl and Ramanathan (1983) is employed in the long-wave parameterization scheme of CAM3. In this approach, an exponential transmission approximation according to Ramanathan et al. (1985) is applied for CO_2, O_3, CH_4 and N_2O to estimate their absorptivity and emissivity. For water vapor, a General Line-by-

line Atmospheric Transmittance and Radiance Model (GLATRM; Collins et al. 2004) is used to calculate the absorptivity and emissivity of water. Clouds are assumed as gray bodies with emissivity defined by cloud phase, condensed water path and effective radius of ice particles.

2.1.2.2.3 Turbulence parameterization

The turbulence parameterization is represented by the estimation of diffusivity for the free atmosphere and by a non-local ABL parameterization that calculates the boundary layer depth. First, the free atmospheric turbulence diffusivity is calculated at all levels. The free atmospheric turbulence diffusivity is represented by the mixing length ($\approx$ 30m), vertical shear and stability function by the gradient Richardson number (R_i) (Collins et al. 2004). The Richardson number is a dimensionless number that relates vertical stability to vertical shear and introduces the condition for stable ($R_i > 0$), neutral ($R_i = 0$), and unstable ($R_i < 0$) stratifications. The stability function is defined by different expressions according to the sign of the Richardson number. In the next step, ABL height is determined from the horizontal wind, virtual potential temperature above the surface layer, a disposable parameter ($\approx$ 100) that determines the turbulence due to surface friction and the critical Richardson number ($\approx$ 0.3) (Vogelzang and Holtslag 1996). Then the free atmosphere values are replaced by the new calculated values for all levels within the ABL.

2.1.3 Common Land Model version 3

CLM3 (Dai et al. 2003, Oleson et al. 2004) is a single column snow-soil-vegetation model that can be run in offline mode or coupled with the atmospheric model (CAM) and CCSM3 through a coupler (cpl). CLM3 is also integrated at the same zonal resolution as CAM3, but CLM3 is designed to use nested sub-grid scale horizontal landscape heterogeneity considered by a mosaic-type approach (e.g. Avissar and Pielke 1989), in which grid-cells are composed of various land units (e.g., glacier, lake, wet land and up to 16 standard Plant Functional Types (PFTs) including "bare" soil), where each grid-cell

can have a different number of land units, each land unit can have snow/soil columns and each column can have multiple PFTs. Biological processes are simulated for each sub-grid land unit, column, and PFT independently and each sub-grid unit maintains its own prognostic variables. CLM3 assumes a single vegetation layer, ten soil layers with 0.007, 0.014, 0.027, 0.052, 0.102, 0.199, 0.388, 0.755, 1.47, and 2.86m depth (with a total thickness 3.43m) and up to five snow layers (depending on snow depth) which represent vertical heterogeneity within the snow-pack.

Vegetation and vegetation-fraction data are obtained from the global 1km×1km International Geosphere-Biosphere Program (IGBP) land cover and 8km×8km green leaf area index data. Soil physical data are from Bonan (1996).

2.1.3.1 Water balance

In the mass conservation equation for water, the water flux is equal to the net flow across the boundary, phase changes, and sources or sinks of water mass. Integration of the mass conservation equation over a control volume of canopy, soil and snow layer under the assumptions of horizontal homogeneity and negligible water vapor results in the water-balance equations for the canopy and soil and snow water, respectively. The time rate of change in actual canopy storage is defined by the sum of precipitation intercepted by canopy, drainage and evaporation at the surface of wet foliage (e.g. Jetten 1996, Dai et al. 2001). The soil-water balance equation considers the difference in water flux through the interface, transpiration and rate of phase change from ice to liquid for liquid water. The soil-ice balance equation considers the difference in ice mass through the interface of two consecutive layers including phase transition changes. The snow-water balance considers the difference in flux and phase transition changes in each snow-model layer. The classical bulk-aerodynamic formulation of water evaporation from the bare soil and snow at reference level, considers air density, aerodynamic resistance and the difference of specific humidity between atmospheric reference level and the surface (Mahfouf and Noilhan 1991, Dai et al. 2001) is used here.

A TOPMODEL-approach (Beven and Kirkby 1979, Campling et al. 2002) in CLM3 (Oleson et al. 2004) considers the topographic control on soil moisture and runoff (Niu et al. 2005). It provides the basis for partitioning a grid-cell into saturated and unsaturated areas to estimate of surface runoff (Wang et al. 2008). Saturated hydraulic conductivity decreases exponentially with height according to the TOPMODEL-approach. The fraction of saturated area is determined from water table depth, and the maximum saturated fraction is estimated by the topographic index distribution, while the unsaturated area is characterized by soil moisture. The total water amount reaching the soil surface is assumed as runoff for an impermeable and permeable topsoil layer; surface runoff is equal to the sum of runoff from saturated and unsaturated areas (Oleson et al. 2004).

The local rate of change of soil water is represented by the soil water flux in the vertical direction according to Darcy's law, phase change from ice to liquid and transpiration (Dai et al. 2001, 2003). The hydraulic conductivity and soil matric potential of soil layers differ with the variation of volumetric soil water and soil texture (Clapp and Hornberger 1978, Cosby et al. 1984). Hydraulic conductivity depends on the water content of two layers and follows the TOPMODEL-approach (Oleson et al. 2004). The soil matric potential depends only on temperature for frozen soils (Fuchs et al. 1978, Oleson et al. 2004). Inclusion of soil water in the model considers the changes in soil-temperature associated with freezing and thawing (Viterbo et al. 1999, Warrach et al. 2001, Mölders and Walsh 2004). However, the interaction between soil-temperature and moisture cross effects like Ludwig-Soret effect and Dufour effect that are important especially in permafrost region (Mölders et al. 2003, Mölders and Walsh 2004), are not considered in this climate model. The uncertainty due to neglecting these cross effects is negligible under most circumstances (Mölders and Walsh 2004).

The local rate of change of snow ice and water at the surface snow layer are represented by solid and liquid forms of precipitation and sublimation and evaporation, respectively (Dai et al. 2001). The water flow in snow layer is a function of porosity and

irreducible water in snow. If the porosity of the snow layer is less than 0.05, water flow within that layer will be assumed to zero (Dai et al. 2001).

The effective root fraction depends on the root fraction within the soil layer and soil matrix potential. The maximum transpiration rate, which is taken from BATS (Dickinson et al. 1993), is a function of fraction of vegetation excluding snow, leaf-area index, root fraction within the soil layer and soil matrix potential (Dai et al. 2001).

2.1.3.2 Energy balance

In the energy balance equation, the rate of change of stored heat depends on convection, conduction and radiation. Generally, three assumptions are considered (Dai et al. 2001):

1. Heat transfer by convection within the canopy, soil and snow is assumed to be zero
2. Conductance of heat within the canopy is zero
3. Vaporization and sensible heat transfer within the snow and soil layer are zero

With these assumptions, the net radiation at canopy can be expressed by the latent heat flux and sensible heat flux, which depend on the canopy temperature (Oleson et al. 2004). Soil- and snow-temperature for ten soil layers with up to five overlaying snow layers are estimated by solving the second law of heat conduction (Oleson et al. 2004). Soil- and snow-temperature are defined by the heat flux, volumetric heat capacity, and latent heat of phase transition. The Crank-Nicholson method serves to solve the heat equation for both snow and soil layers (Dai et al. 2003, Oleson et al. 2004).

2.1.3.3 Surface albedo

CLM3 considers three types of ground albedo: snow, soil and vegetation albedo for visible ($<0.7\mu m$) and near-infrared ($\geq 0.7\mu m$) solar radiation. Soil and snow albedo stem from the Biosphere-Atmosphere Transfer Scheme (BATS, Dickinson et al. 1993). Visible and near-infrared snow albedo for diffuse incident radiation is a function of snow age

(fractional decrease of albedo due to snow aging, i.e., increasing grain size, dirt and soot content); albedo for visible and near-infrared solar radiation incident on fresh snow with a solar zenith angle less the 60^{o} and snow albedo for direct beam depend on albedo for diffuse beam and the cosine of the solar zenith angle. Surface albedo differs for glaciers, lakes and wetland and soil surfaces. Glacier, frozen lake and wet land albedo are adopted from the NCAR Land Surface Model (LSM, Bonan 1998). Albedo of unfrozen lake and wet land are functions of the cosine of the solar zenith angle (Oleson et al. 2004). Bare soil albedo for visible radiation is represented by saturated soil albedo and the change in albedo due to dryness and soil surface color. Snow and soil albedo for direct incident radiation are assumed to be nearly equal to the snow albedo for diffuse incident radiation (Dai et al. 2001).

Vegetation albedo is adopted from the LSM to capture the essential features of the two-stream approximation model. Vegetation albedo for the direct beam is represented by summing up the albedo of the underlying surface and ground reflection. The parameterization of vegetation albedo for the direct beam neglects multiple reflections between canopy and underlying ground surface (Dai et al. 2001). If the cosine of the solar zenith angle equals 0.5, vegetation albedo for the direct beam will equal the vegetation albedo for the diffuse incident radiation beam.

2.1.3.4 Radiative flux

The net solar flux absorbed by the ground is the difference between the total solar flux absorbed by the surface and vegetation. The net solar flux absorbed by the surface is a function of incident visible and near-infrared solar radiation and surface albedo. The net solar flux absorbed by vegetation is represented by incident visible and near-infrared solar radiation, the fraction of solar radiation absorbed by vegetation and the snow-free vegetation fraction (Dai et al. 2001). The net long-wave radiation absorbed at the ground and vegetation are functions of the snow-free vegetation fraction, emissivity of the vegetation and ground, and downward long-wave radiation (Dai et al. 2001).

2.1.3.5 Turbulent flux

Aerodynamic roughness and zero-plane displacement height are defined as a function of canopy height, vegetation density, leaf-area index and leaf shape. The zonal and meridional momentum flux, the sensible heat flux and water vapor flux between the atmosphere at reference height and canopy top or bare ground are calculated based on the Monin-Obukhov similarity hypothesis (Dai et al. 2003). According to this hypothesis, dimensionless mean horizontal wind speed, mean potential temperature, and mean specific humidity logarithmic profiles depend on unique stability functions (Zeng et al. 1998, Oleson et al. 2004). This stability parameter is represented by the ratio of the difference between surface layer height and the displacement height, to the Monin-Obukhov length. The Monin-Obukhov length is a function of acceleration due to gravity, the von Kárman constant, virtual potential temperature, turbulent scale-temperature and friction velocity at the surface. Positive and negative Monin-Obukhov lengths describe stable and unstable conditions, respectively. For neutral conditions, the Monin-Obukhov length tends to infinity and hence the stability function becomes zero.

The momentum flux, sensible heat flux and water vapor fluxes between surface and reference height are represented by air density, aerodynamic resistance for momentum, sensible heat and water vapor, wind speed at the reference height, the difference between the lowest level and surface specific humidity, and the difference between reference height and surface potential temperatures.

2.1.4 Community Sea Ice Model version 5

Community Sea Ice Model version 5 (CSIM5; Briegleb et al. 2004) is designed to run with approximately 1^{o} horizontal resolution with 320 zonal points, 384 meridional points and 40 layers in vertical direction. CSIM5 considers the elastic-viscous-plastic dynamics scheme (Hunke and Dukowicz 1997), energy conserving thermodynamics (Bitz and Lipscomb 1999), Lagrangian distribution of ice thickness (Thorndike et al. 1975, Bitz et al. 2001), a remapping scheme (Lipscomb and Hunke, 2004) for estimating horizontal advection, and lateral and bottom melt processes (McPhee 1992).

2.1.5 Parallel Ocean Program version 1.4.3

Parallel Ocean Program (POP) version 1.4.3 (Smith and Gent 2004) has the same resolution as the sea ice component of CCSM3. In POP, ocean dynamics are treated by the 3D primitive equations for a stratified fluid in the general orthogonal coordinate with hydrostatic and Boussinesq approximation, and by the barotropic equation with a linearized implicit free-surface scheme. POP uses the transport parameterization on constant potential density surfaces with a mixing coefficient of $600m^2/s$ (Large et al. 2001), a K-profile parameterization for vertical mixing (Large et al. 1994), anisotropic horizontal viscosity (Smith and McWilliams 2003) and has a diurnal cycle of solar forcing. In this component, the North Pole is displaced to Greenland (Holland 2003).

References

Avissar R, Pielke RA (1989) A parameterization of heterogeneous land surfaces for atmospheric numerical models and its impact on regional meteorology. Mon Wea Rev 117: 2113-2136

Berger AL (1978) Long-term variations of daily insolation and quaternary climatic changes. J Atmos Sci 35: 2362-2367

Beven KJ, Kirkby MJ (1979) A physically based, variable contributing area model of basin hydrology. Hydrol Sci Bull 24: 43-69

Bitz CM, Lipscomb WH (1999) An energy-conserving thermodynamic model of sea ice. J Geophys Res 104: 15,669-15,677

Bitz CM, Holland M, Eby M, Weaver AJ (2001) Simulating the ice-thickness distribution in a coupled climate model. J Geophys Res 106: 2441–2463

Bonan GB (1996) A land surface model (LSM version 1.0) for ecological, hydrological, and atmospheric studies: Technical description and user's guide. NCAR Technical Note NCAR/TN-417+STR, National Center for Atmospheric Research, Boulder, CO

Bonan GB (1998) The land surface climatology of the NCAR Land Surface Model coupled to the NCAR Community Climate Model. J Climate 11: 1307-1326

Briegleb BP, Hunke EC, Bitz CM, Lipscomb WH, Holland MM, Schramm JL, Moritz RE (2004) The sea ice simulation of the Community Climate System Model, version two. NCAR Tech Note NCAR TN-45+STR, 34pp

Campling P, Gobin A, Beven K, Feyen J (2002) Rainfall-runoff modeling of a humid tropical catchment: the TOPMODEL approach. Hydrol process 16: 231-253

Chen C, Cotton WR (1987) The physics of the marine stratocumulus-capped mixed layer. J Atmos Sci 44: 2951-2977

Clapp RB, Hornberger M (1978) Empirical equation for some soil hydraulic properties. Water Resour Res 14: 601-604

Collins WD (2001) Parameterization of generalized cloud overlap for radiative calculations in general circulation models. J Atmos Res 58: 3224-3242

Collins WD, Rasch PJ, Boville BA, Hack JJ, McCaa JR, Williamson DL, Kiehl JT, Briegleb BP, Bitz C, Lin SJ, Zhang M, Dia Y (2004) Description of the NCAR Community Atmospheric Model (CAM3). Technical report NCAR/TN464+STR, National Center for Atmospheric Research, Boulder, CO, pp 226

Collins WD, Bitz CM, Blackmon ML, Bonan GB, Bretherton CS, Carton JA, Chang P, Doney SC, Hack JJ, Henderson TB, Kiehl JT, Large WG, McKenna DS, Santer BD, Smith RD (2006a) The Community Climate System Model: CCSM3. J Climate 19: 2122-2143

Collins WD, Rasch PJ, Boville BA, Hack JJ, McCaa JR, Williamson DL, Briegleb BP, Bitz CM, Lin SJ, Zhang M (2006b) The formulation and atmospheric simulation of the Community Atmosphere Model: CAM3. J Climate 19: 2144-2161

Cosby BJ, Hornberger GM, Clapp RB, Ginn TR (1984) A statistical exploration of the relationship of soil moisture charteristics to the physical properties of soils. Water Resour Res 20/6:682-690

Dai Y, Zeng X, Dickinson RE (2001) Common Land Model (CLM) (Available online at http://climate.eas.gatech.edu/dai/clmdoc.pdf 2001)

Dai Y, Zeng X, Dickenson RE, Backer I, Bonan GB, Bosilovich MG, Denning AS, Dirmeyer PA, Houser PR, Niu G, Oleson KW, Schlosser CA, Yang ZL (2003) The Common Land Model. Bull Amer Meteor Soc 84: 1013-1023

Dickinson RE, Henderson-Sellers A, Kennedy PJ, Wilson MF (1993) Biosphere atmosphere transfer scheme (BATS) version 1e as coupled for Community Climate Model. NCAR Tech Note NCAR/TN-378+STR, 72pp

Dickinson RE, Oleson KW, Bonan G, Hoffman F, Thornton P, Vertenstein M, Yang Z-L, Zeng X (2006) The Community Land Model and its climate statistics as a component of the Community Climate System Model. J Climate 19: 2302-2324

Fuchs M, Campbell GS, Papendick RI (1978) An analysis of sensible and latent heat flow in a partially frozen unsaturated soil. Soil Sci Soc Am J 42: 379-385

Hack JJ, Boville BA, Briegleb BP, Kiehl JT, Rasch PJ, Williamson DL (1993) Description of the NCAR Community Climate Model (CCM2), Technical Report NCAR/TN-382+STR, 120 pp

Hack J (1994) Parameterization of moist convection in the National Center for Atmospheric Research community climate model (CCM2). J Geophys Res 99: 5551-5568

Holland M (2003) The North Atlantic oscillation-Arctic oscillation in the CCSM2 and its influence on Arctic climate variability. J Climate 16: 2767-2781

Hunke EC, Dukowicz JK (1997) An elastic-viscous-plastic-model for sea ice dynamics. J Phys Oceanogr 27: 1849-1867

Jetten VG (1996) Interception of Tropical rain forest: performance of a canopy water balance model. Hydrol Process 10: 671-685

Kessler E (1969) On the Distribution and Continuity of Water Substance in Atmospheric Circulations. Amer Meteor Soc, Boston, 84pp

Kiehl JT, Ramanathan V (1983) CO_2 radiative parameterization used in climate models: comparison with narrow band models and with laboratory data. J Geophys Res 88: 5191-5202

Kiehl JT, Ramanathan V (1990) Comparison of cloud forcing derived from the earth radiation budget experiment with that simulated by the NCAR community climate model. J Geophys Res 95: 11,679-11,698

Kiehl JT, Briegleb B (1991) A new parameterization of the absorbtance due to the 15μm band system of carbon dioxide. J Geophys Res 96: 9013-9019

Kiehl JT, Hack JJ, Bonan GB, Boville BB, Williamson DL, Rasch PJ (1998) The National Center for Atmospheric Research Community Climate Model: CCM3. J Climate 11: 1131-1149

Large WG, McWilliams JC, Doney SC (1994) Oceanic vertical mixing: A review and a model with a non-local boundary layer parametrization. Rev Geophys 32: 363-403

Large WG, Danabasoglu G, McWilliams JC, Gent P, Bryan F (2001) Equatorial circulation of a global ocean climate model with anisotropic horizontal viscosity. J Phys Oceanogr 31:518-536

Lin YL, Farley RR, Orville HD (1983) Bulk parameterization of the snow field in a cloud model. J Climmate Appl Meteor 22: 1065-1092

Lipscomb WH, Hunke EC (2004) Modeling sea ice transport using incremental remapping. Mon Wea Rev 130: 1848-1865

Mahfouf JF, Noilhan J (1991) Comparative study of a various formulation of evaporation from bare soil using situ data. J Appl Meteor 30: 1354-1365

McPhee MG (1992) Turbulent heat flux in the upper ocean under sea ice. J Geophys Res 97: 5365-5379

Mölders N, Haferkorn U, Döring J, Kramm G (2003) Long-term numerical investigations on the water budget quantities predicted by the hydro-thermodynamic soil vegetation scheme (HTSVS) - Part I: Description of the model and impact of long-wave radiation, roots, snow, and soil frost. Meteor Atmos Phys 84: 115-135

Mölders N, Walsh JE (2004) Atmospheric response to soil-frost and snow in Alaska in March. Theor Appl Climatol 77: 77-105

Niu GY, Yang ZL, Dickinson RE, Gulden LE (2005) A simple TOPMODEL-based runoff parameterization (SIMTOP) for use in GCMs. J Geophys Res 110: D21106 doi: 10.1029/2005JD006111

Oleson KW, Dai Y, Bonan GB, Bosilovich M, Dickinson R, Dirmeyer P, Hoffman F, Houser P, Levis S, Niu GY, Thornton P, Vertenstein M, Yang ZL, Zeng X (2004) Technical description of Community Land Model (CLM). Technical Report NCAR/TN-461+STR, National Center for Atmospheric Research, Boulder, CO. 80307-3000, pp 174

Ramanathan V (1976) Radiative transfer within the Earth's troposphere and stratosphere: A simplified radiative-convective model. J Atmos Sci 33: 1330-1346

Ramanathan V, Cicerone RJ, Singh HB, Kiehl JT (1985) Trace gas trends and their potential role in climate change. J Geophys Res 90: 5547-5566

Rasch PJ, Mahowald NM, Eaton BE (1997) Representations of transport, convection, and the hydrologic cycle in chemical transport models: Implications for the modeling of short-lived and soluble species. J Geophys Res 102: 28127-28138

Rasch PJ, Kristjánsson JE (1998) A comparison of the CCM3 model climate using diagnosed and predicted condensate parameterizations. J Climate 11: 1587-1614

Slingo JM (1987) The development and verification of a cloud prediction scheme for the ECMWF model. Quart J Roy Meteor Soc 113: 899-927

Smith RD, Dukowicz JK, Malone RC (1992) Parallel ocean general circulation modeling. Physica D 60: 38-61

Smith RD, McWilliams JC (2003) Anisotropic horizontal viscosity for ocean model. Ocean Modell 5: 129-156

Smith RD, Gent P (2004) Anisotropic GM parameterization of ocean model. J Phys Ocean 34:2541-2564

Stowe LL, Ignatov AM, Singh RR (1997) Development, validation, and potential enhancements to the second-generation operational aerosol product at the National Environmental Satellite, Data, and Information Service of the National Oceanic and Atmospheric Administration. J Geophys Res 102: 16889-16910

Sundqvist H (1978) A parameterization scheme for non-convective condensation including prediction of cloud water content. Quart J Roy Meteor Soc 104: 677-690

Sundqvist H (1988) Parameterization of condensation and associated clouds in models for weather prediction and general circulation simulation. In Schlesinger ME (ed), Physically-based modeling and simulation of climate and climate change. Vol. 1, Kluwer Academic Publishers, Dordrecht: 433-461

Thorndike AS, Rothrock DS, Maykut GA, Colony R (1975) The thickness distribution of sea ice. J Geophys Res 80: 4501–4513

Tripoli GJ, Cotton WR (1980) A numerical investigation of several factors contributing to the observed variable intensity of deep convection over south Florida. J Appl Meteor 19: 1037-1063

Twomey SA, Piepgrass M, Wolfe TL (1984) An assessment of the impact of pollution on global cloud albedo. Tellus 36: 356–366

Viterbo P, Beljaars A, Mahouf JF, Teixeira J (1999) The representation of soil moisture freezing and its impact on the stable boundary layer. Quart J Roy Meteor Soc 125: 2401-2447

Vogelzang DHP, Holtslag AAM (1996) Evaluation and model impacts of alternative boundary-layer height formulations. Bound-Layer Meteor 81: 245-269

Wang A, Li KY, Lettenmaier DP (2008) Integration of the variable infiltration capacity model soil hydrology scheme into the community land model. J Geophys Res 113, D09111, doi:10.1029/2007JD009246

Warrach K, Mengelkamp HT, Raschke E (2001) Treatment of frozen soil and snow cover in the land surface model SEWAB. Theor Appl Climatol 69: 23-37

Yanai M, Esbensen S, Chu JH (1973) Determination of bulk properties of tropical cloud clusters from large-scale heat and moisture budgets. J Atmos Sci 30: 611-627

Zeng X, Dickinson RE, Shaikh M (1998) Effect of surface sublayer on surface skin temperature and fluxes. J Climate 11: 537-550

Zhang GJ, McFarlane NA (1995) Sensitivity of climate simulations to the parameterization of cumulus convection in the Canadian Climate Centre general circulation model. Atmos-Ocean 33: 407- 446

Zhang M, Lin W, Bretherton CS, Hack JJ, Rasch PJ (2003) A modified formulation of fractional stratiform condensation rate in the NCAR community atmospheric model CAM2. J Geophys Res 108D 4035 doi: 10.1029/2002JD002523

Chapter 3 Methodology

3.1 CCSM3 simulations

The fully-coupled CCSM3 is run with 26 vertical layers at a spatial resolution of $\approx 2.8^{o}$x2.8^{o} (T42). As pointed out above, the reference simulation is performed for 50 years and started with the ecliptic conditions of 1-1-1950. It assumes a constant CO_2 concentration of 335 ppmv for the 50 years reference run. Each model component of CCSM3 is spun up separately.

The CCSM3 simulation for 1950-2000 is considered as the reference simulation for this study. Based on these 50 years of simulations, three climatologies 1951-1980, 1961-1990, and 1971-2000 are determined. Soil-temperature climatologies obtained from the reference CCSM3 simulation are compared with observed climatologies over Russia to evaluate CCSM3's performance and determine the sources of the errors. Uncertainties in soil parameters may introduce some errors in simulated soil-temperatures especially during the freezing of soil water and thawing of frozen ground (Mölders et al. 2003, 2005). To investigate the influence of soil-parameter inaccuracy on simulated soil-temperatures, two CCSM3 simulations wherein the sand fraction is reduced (enhanced) by 10 % and the clay fraction is increased (decreased) by 10 % in Russia are performed for 1950-2000. These simulations are denoted S10D and S10I, respectively. The descriptions of different simulations are discussed in detail in Chapter 4

3.2 Observational data

CCSM3 soil climatologic performance is evaluated using observed soil-temperature data from National Snow Ice Data Centre (NSIDC; Zhang et al. 2001) at 0.2, 0.4, 0.8, 1.6, and 3.2 m depth. Observations at 0.8 m and deeper were made daily near midday, while those at shallower depths were made at 0100, 0700, 1300, and 1900 UT until 1965, and every 3 hours starting at 0000 UT later on.

European Centre for Medium-Range Weather Forecasts (ECMWF) 40 years reanalysis (ERA40) near-surface air temperatures data (Uppala et al. 2005) with 2.5^{o}x2.5^{o}

resolution are used to evaluate the influence of near-surface air temperatures biases on soil-temperature simulations. The monthly precipitation data is obtained from the Global Precipitation Climatology Centre (GPCC) at 2.5°x2.5° resolution for 1951-2000. The observed cloud-cover-fraction data is available for 1984-2004 on a 2.5°x2.5° grid from the International Satellite Cloud Climatology project (ISCCP; Rossow and Schiffer 1999). ISCCP total cloud-cover fraction is represented by the percentage of pixels identified as cloudy within a grid-cell. After interpolating these observed data to the 2.8°x2.8° CCSM3 grid, climatologies for near-surface air temperature, precipitation and cloud-cover fraction are calculated and compared with their simulated values over Russia to assess the impact of model biases on soil-temperature simulations. Snow-depth data is available for 1978-2003 from in-situ measurements at over 3000 stations and provided by Zhang (2006 pers. Communication). The descriptions of observational data are discussed in detail in Chapter 4.

3.3 WRF simulations

To assess the impact of network density and design on regional averages (see the question c in the introduction), Weather Research and Forecasting (WRF; Skamarock et al. 2005) model simulations are performed over Russia for July and December 2005, 2006, 2007 with 70×150 grid points, a 50 km grid-increment, 30 vertical layers from the surface to 50 hPa and six layers within the soil. More details on the model set up are given in chapter 5. The WRF model is run for July and December 2005, 2006 and 2007. These months are chosen as CCSM3-simulated and gridded soil-temperature data differed the most in these months. Years outside the 50 years, for which the CCSM3 is run, are chosen to have an independent dataset.

The WRF simulations are performed for three July and December months to enhance the number of possible weather situations. The WRF simulations serve to create a dataset of soil-temperature and atmospheric conditions with a high resolution. The advantage of using WRF simulations results over artificial data is that it permits also to consider the aspects related to landscape and latitude.

Regional averages for various quantities for area $2.8^{o} \times 2.8^{o}$ determined from all WRF-output are considered as "reference". Regional averages for $2.8^{o} \times 2.8^{o}$ calculated based on forty artificial networks consisting of ten sets of 500, 400, 200, and 100 randomly chosen WRF-grid-points as "sites" and 411 "sites" that correspond to the locations of a real or historic network, are compared with the reference regional averages. The comparison of regional averages derived from various networks and the reference network reveals that networks with 200 or more randomly distributed sites reliably reproduce the reference regional averages while the real network has difficulties in capturing the reference regional averages due to its non-random distribution of sites. The results of WRF investigations are described in detail in Chapter 5 and 6.

3.4 Analysis

CCSM3's LSM reasonably well simulates land-surface processes, snow accumulation and snow-melt, state variables, and fluxes in offline mode (Bonan et al. 2002, Dai et al. 2003, Bonan and Levis 2006, Lannoy et al. 2006). Nicolsky et al. (2007) evaluated the soil-temperature simulated by CLM3 in its offline mode by means of in-situ data over Alaska.

To evaluate the fully-coupled soil-temperature simulations and determine sources of error, soil-temperature climatologies simulated by CCSM3 are compared with climatologies derived from observations in Russia provided by NSIDC (Zhang 2006 pers. Communications) in this dissertation. To test the research hypothesis and answer the questions brought up in the introduction the following steps are taken.

CCSM3-simulated soil-temperatures are compared with gridded soil-temperatures derived from observation. Systematic and random errors can contribute to any simulation error. To evaluate performance and attribute error sources at different spatial and temporal scales, performance measures (Anthes 1983, Anthes et al. 1989) for the annual course and mean for all climatologies for both domain-averages and each grid-cell are calculated. This analysis answers the question how well CCSM3 perform in a fully coupled mode.

Since near-surface air temperature, cloud fraction, precipitation and snow depth can drive soil-temperature, these quantities simulated by CCSM3 are also evaluated with respect to ERA40 reanalysis, ISCCP, GPCC and NSIDC snow depth data, respectively. The performance measures are also determined between CCSM3-simulated quantities and their observations.

To test the hypothesis that incorrectly-simulated precipitation, snow depth, and/or cloud fraction cause errors in simulated soil-temperatures we apply a Student's t-test to the differences between CCSM3-simulated and observed forcing quantities and the differences between CCSM3-simulated and observed soil-temperatures. Note that in the following the word *significant* will only be used if differences or correlation pass a t-test or R-test at the 95% or higher confidence level. This analysis provides the first answer to the question on the reasons for the discrepancies between CCSM3-simulated and gridded soil-temperatures.

To assess the impact of network density and design on regional averages, performance measures are also determined between the reference regional averages and regional averages determined from the various artificial networks and the real network to investigate the performance of a network in reproducing the reference regional average. To measure the strength of the various networks in capturing trends and/or phases of regional averages and determine the overall relative degree to which the regional averages derived from various networks approach the reference regional averages, correlation-skill scores and Willmott's index of agreement (Willmott 1984) between the regional averages derived for the various networks and the reference regional averages are determined, respectively. Willmott's index of agreement (WIA) ranges from 0 (complete disagreement) to 1 (perfect agreement). This analysis provides the answers to the question how the network density and/or design affects gridded data and hence affects the discrepancies between CCSM3-simulated and gridded soil-temperatures. The detail descriptions of analysis methods are discussed in Chapter 4, 5 and 6.

References

Anthes RA (1983) Regional models of the atmosphere in middle latitudes. Mon Wea Rev 111: 1306-1335

Anthes RA, Kuo YH, Hsie EY, Low-Nam S, Bettge TW (1989) Estimation of skill and uncertainty in regional numerical models. Quart J Roy Meteorol Soc 111: 763-806

Bonan GB, Oleson KW, Vertenstein M, Levis S, Zeng X, Dai Y, Dickinson RE, Yang Z-L (2002) The Land Surface Climatology of the Community Land Model Coupled to the NCAR Community Climate Model. J Climate 15: 3123-3149

Bonan GB, Levis S (2006) Evaluating aspects of the Community Land and Atmosphere Models (CLM3 and CAM3) using the CLM's dynamic global vegetation model. J Climate 19: 2290-2301

Dai Y, Zeng X, Dickinson RE, Baker I, Bonan G, Bosilovich M, Denning S, Dirmeyer P, Houser P, Niu G, Oleson K, Schlosser A, Yang ZL (2003) The Common Land Model (CLM). Bull Amer Meteor Soc 84: 1013-1023

Lannoy GD, Houser PR, Pauwels VRN (2006) Assessment of model uncertainty for soil moisture through ensemble verification. J Geophys Res 111, D10101, doi:10.1029/2005JD006367

Mölders N, Haferkorn U, Döring J, Kramm G (2003) Long-term numerical investigation on the water budget quantities predicted by the hydro-thermodynamic soil vegetation scheme (HTSVS) – Part I: Description of the model and impact of long-wave radiation, roots, snow, and soil frost. Meteor Atmos Phys 84: 115-135

Mölders N, Jankov M, Kramm G (2005) Application of Gaussian error propagation principles for theoretical assessment of model uncertainty in simulated soil processes caused by thermal and hydraulic parameters. J. Hydrometeor 6: 1045-1062

Nicolsky DJ, Romanovsky VE, Alexeev VA, Lawrence DM (2007) Improved modeling of permafrost dynamics in a GCM land-surface scheme. Geophys Res Lett 34: L08501, doi:10.1029/2007GL029525

Rossow WB, Schiffer RA (1999) Advances in understanding clouds from ISCCP. Bull Amer Meteor Soc 80: 2261-2288

Skamarock WC, Klemp JB, Dudhia J, Gill DO, Baker DM, Wang W, Powers JG (2005) A description of the advanced research WRF version 2. NCAR Technical Note, NCAR/TN-468+STR, pp. 88

Uppala SM, Kållberg PW, Simmons AJ, Andrae U, Bechtold VD, Fiorino M, Gibson JK, Haseler J, Hernandez A, Kelly GA, Li X, Onogi K, Saarinen S, Sokka N, Allan RP, Andersson E, Arpe K, Balmaseda MA, Beljaars ACM, Berg L, Bidlot J, Bormann N, Caires S, Chevallier F, Dethof A, Dragosavac M, Fisher M, Fuentes M, Hagemann S, Hólm E, Hosking BJ, Isaksen L, Janssen PAEM, Jenne R, McNally AP, Mahfouf JF, Morcrette JJ, Rayner NA, Saunders RW, Simon P, Sterl A, Trenberth KE, Untch A, Vasiljevic D, Viterbo P, Woollen J (2005) The ERA-40 reanalysis. Q J R Meteorol Soc 131: 2961-3012

Willmott CJ (1984) On the evaluation of model performance in physical geography. In: G.L. Gaile and C.J. Willmott (eds). Spatial Statistics and Models. Dordrecht, Holland: D. Reidel: 443-460

Zhang T, Barry RG, Gilichinsky D, Bykhovets SS, Sorokovikov VA, Ye J (2001) An amplified signal of climatic change in soil temperatures during the last century at Irkutsk, Russia. Climatic Change 49: 41-76

Chapter 4 Evaluation of Community Climate System Model soil temperatures using observations from Russia*

Abstract

Soil temperatures simulated by the fully coupled Community Climate System Model (CCSM) version 3.0 are evaluated using three gridded climatologies (1951-1980, 1961-1990, 1971-2000) based on data from more than 400 Russian sites. CCSM captures the annual phase of the soil temperature cycle well, but not the amplitude. It provides slightly too high (low) soil temperatures in winter (summer). Root mean square errors, on average, are less than 5K.

Simulated near-surface air temperatures agree well, on average, with near-surface air temperatures from reanalysis data. Errors in simulated atmospheric-temperature forcing correlate statistically significantly (95% or higher confidence level) with soil temperature errors, i.e. contribute to discrepancy in soil temperature simulation. Comparison to International Satellite Cloud Climatology project data shows that errors in simulated cloud fraction explain some soil and near-surface air temperature and precipitation discrepancies. Evaluation by means of Global Precipitation Climatology Centre data identifies inaccurately-simulated precipitation as a contributor to underestimating summer soil temperatures. Comparison to snow-depth observations shows that overestimating snow depth yields winter soil-temperature overestimation.

Sensitivity studies show that uncertainty in mineral-soil composition notably, and differences between the vegetation in CCSM and nature marginally contribute to discrepancies between simulated and observed soil-temperature climatology.

---

* PaiMazumder D, Miller J, Li Z, Walsh JE, Etringer A, McCreight J, Zhang T, Mölders N (2008) Evaluation of Community Climate System Model soil temperatures using observations from Russia. Theor Appl Climatol 94: 187-213

4.1 Introduction

Permafrost (soil or rock that remains at or below 0°C for at least two consecutive years) occupies about $22.79 \cdot 10^6$ km^2 (23.9 %) of Northern Hemispheric land with approximately 70 % occurring between 45°N and 67°N (Zhang et al. 1999). Soil above the permafrost that annually freezes and thaws is called the active layer. Soil temperature is primarily controlled by the surface water and energy balance, which explains the strong linkage and feedback between soil and near-surface atmospheric conditions. Among other things, Arctic and Subarctic soil temperatures strongly depend on conditions of seasonal snow-cover, such as onset, duration, thickness, density and structure (e.g. Zhang et al. 1996, Mölders and Romanovsky 2006).

Permafrost temperature and active layer depth are sensitive to climatic variability (e.g. Kane et al. 1991, Osterkamp 2003); concurrently, thermal and hydrological conditions associated with permafrost and the active layer affect climate by heat, moisture, and trace gas exchange (e.g. Stendel and Christensen 2002, Mölders and Walsh 2004). Understanding feedbacks between permafrost, weather and climate, and other potential impacts like economic and infrastructure damages, ecosystem changes, and freshwater availability (Esch and Osterkamp 1990, Cherkauer and Lettenmaier 1999, Oechel et al. 2000, Serreze et al. 2000, Zhuang et al. 2001) requires adequate soil-temperature simulations in numerical weather prediction models (NWPMs), General Circulation Models (GCMs), and Earth System Models (ESMs). Therefore, great efforts have been made to implement frozen-soil physics into the land surface models (LSMs) of NWPMs, GCMs and ESMs (e.g. Koren et al. 1999, Viterbo et al. 1999, Boone et al. 2000, Warrach et al. 2001, Bonan et al. 2002, Dai et al. 2003, Mölders et al. 2003, Bonan and Lewis 2006).

To better understand land-surface processes and differences in LSM performance the Project for Intercomparison of Land Surface Parameterization Schemes (PILPS) compared off-line simulations of various LSMs; the main differences found were spin-up time and flux partitioning (e.g. Henderson-Sellers et al. 1993, 1995). PILPS' second phase evaluated LSM offline performance using observations (e.g. Schlosser et al. 2000,

Slater et al. 2001, Luo et al. 2003). A main finding was that average annual water and energy partitioning appreciably differed among schemes. PILPS studies focusing on LSM performance at high latitudes showed that peat-lands and bogs introduce errors (Bowling et al. 2003). Due to LSM complexities, it is difficult to generalize causes for differences (Nijssen et al. 2001). However, inclusion of frozen-ground physics was identified as important for simulating energy and water fluxes at the earth's surface (e.g. Cherkauer and Lettenmaier 1999) and soil temperature and soil moisture states (e.g. Montaldo and Albertson 2001). Capturing freezing front depth is also important, because temperature variations diminish greatly in the deeper soil (e.g. Luo et al. 2003, Mölders and Romanovsky 2006).

Evaluating simulations of LSMs run in coupled mode with a NWPM, ESM, or GCM is a high priority of the third PILPS phase (e.g. Henderson-Sellers et al. 1995), because errors in simulated atmospheric forcing can affect near-surface conditions with further feedback to simulated soil conditions (e.g. Henderson-Sellers et al. 1995, Chen and Dudhia 2001, Zeng et al. 2002, Narapusetty and Mölders 2005). Choice of initial conditions, discretization, grid resolution, numerical scheme, parameterizations, model assumptions, and/or empirical parameters can lead to incorrectly simulated atmospheric forcing (e.g. Anthes et al. 1989, Slater et al. 1998, Narapusetty and Mölders 2005); soil-temperature simulations themselves can be incorrect for the same reasons. Results obtained from LSMs coupled to GCMs, for instance, differ on the same order of magnitude as off-line PILPS experiments (e.g. Henderson-Sellers et al. 1995); results gained from standard bucket-type hydrological and bio-physically based LSMs, each coupled with GCMs, manifest that LSM complexity may cause statistically significant differences in temperature, pressure, and turbulent fluxes over land (e.g. Sato et al. 1989, Thompson and Pollard 1995, Yang et al. 1995, Qu and Henderson-Sellers 1998). Therefore, even though a LSM is able to capture soil-temperature conditions well when run in offline mode with known site-specific parameters and observed atmospheric forcing, it must be re-evaluated when run in fully-coupled mode with a NWPM, GCM or ESM (e.g. Henderson-Sellers et al. 1995, Chen and Dudhia 2001). Therefore, evaluation

of LSMs in offline and coupled modes must be considered as a necessary and sufficient condition for evaluation, respectively.

In the present study, we evaluate soil temperatures simulated by the Community Climate System Model (CCSM) version 3 (Collins et al. 2006a) using 50 years of Russian soil-temperature measurements. We use observed snow-depth, cloud and precipitation climatologies, and reanalysis air-temperature data plus CCSM sensitivity studies to identify reasons for discrepancies between simulated and observed soil temperatures, and to distinguish between errors resulting from the coupling and the soil model itself.

4.2 Experimental design

4.2.1 Brief model description

CCSM is a fully-coupled climate-system model consisting of the Climate Atmosphere Model (CAM) version 3 (Collins et al. 2006b), the Common Land Model (CLM) version 3 (Dai et al. 2003, Oleson et al. 2004, Dickinson et al. 2006), the Community Sea Ice Model (CSIM) version 5 (Briegleb et al. 2004) and the Parallel Ocean Program (POP) version 1.4.3 (Smith et al. 1992). These four components exchange data via coupler, without flux correction (e.g. Blackmon et al. 2001, Kiehl et al. 2006).

4.2.1.1 Clouds and precipitation

Moist processes are considered by a deep convection scheme (Zhang and McFarlane 1995), shallow convection scheme (Hack 1994) and bulk-microphysics parameterization (Rasch and Kristjánsson 1998, Zhang et al. 2003). The deep convective scheme considers updraft and downdraft and relates cloud-base mass flux to the Convective Available Potential Energy (Collins et al. 2004). Deep convection interacts with large-scale dynamics through pressure-field perturbations caused by cloud-momentum transport (Zhang and McFarlane 1995, Collins et al. 2004). The bulk-microphysical parameterization considers conversion from condensate to precipitating particles, and condensate evaporation, depending on relative humidity (Sundquist 1988, Rasch and

Kristjánsson 1998). At 100 % relative humidity, large-scale stratiform clouds are assumed that depend on moisture and heat advection and cloud water tendencies. For relative humidity from 75-100 %, sub-grid-scale non-convective clouds are considered that depend on moisture and heat advection, cloud water import, and condensation and evaporation of rain or snow (Collins et al. 2004). Clouds in adjacent layers maximally overlap; clouds in layers separated by cloud-free layers randomly overlap (Dai et al. 2003).

4.2.1.2 Land-surface processes

Sub-grid-scale landscape heterogeneity is considered by a mosaic-approach, in which up to four land-cover types (glacier, lake, wetland, vegetation) are distinguished; the vegetation tile is further divided into dominant and secondary plant function type (PFT). Vegetation and vegetation-fraction data stem from the global 1 km x 1 km International Geosphere and Biosphere Program (IGBP) land cover and 8 km x 8 km green leaf area index data. Soil physical data are from Bonan (1996).

Prognostic variables are canopy temperature, water amount in canopy interception storage, snow thickness, and temperature, water, and ice amount within each soil- or snow-model layer. Since according to theory soil-temperature simulations are optimal on a logarithmic grid (e.g. Pielke 2002), we set the soil levels at which to simulate these quantities at 0.007, 0.014, 0.027, 0.052, 0.102, 0.199, 0.388, 0.755, 1.47, and 2.86 m depth. This choice places the soil model bottom at 4 m depth.

The soil-water balance equation considers freezing/thawing of soil water in accord with Fuchs et al. (1978), soil-water uptake by roots, and Darcy's law. Snow temperatures and snow-water equivalent are simulated in up to five layers depending on snow depth. Soil- and snow-temperature calculations consider heat released or consumed by freezing or melting. Energy transfer considers conduction and is solved by a Crank-Nicholson scheme for both soil and snow layers. Convective heat transfer within the canopy, snow, and soil layers, and latent and sensible heat transfer within snow and soil layers are

assumed to be negligible. A TOPMODEL approach accounts for topographic control on soil water and runoff generation (e.g. Niu et al. 2005).

4.2.2 Observations

Soil-temperature data are available from about 400 Russian stations. Soil temperature was measured at 0.2, 0.4, 0.8, 1.6, and 3.2 m depth by extraction thermometers enclosed in an ebonite pipe. The accuracy of these high-inertia mercury thermometers is ±0.1 K (Gilichinsky et al. 1998, Zhang et al. 2001, Romanovsky et al. 2007). Observations at 0.8 m and deeper were made daily near midday, while those at shallower depths were made at 0100, 0700, 1300, and 1900 UT until 1965, and every 3 hours starting at 0000 UT later on. Measurements were made underneath a grass plot and, when snow was present, with natural snow-cover. Due to heat conduction by the pipe soil-temperature measurements may be slightly ($\leq$0.5 K) biased toward colder (warmer) temperatures in the uppermost levels in winter (summer) (Romanovsky 2006; pers. communication).

Comparison between soil temperature simulated for a 2.8°x2.8°-grid-cell and a point measurement are difficult from a scale point of view (e.g. Grayson and Western 1998, Friedrich and Mölders 2000, González-Rouco et al. 2003). Simulated soil temperatures are volume-averages representing model-layer thickness times grid-cell area. Various procedures have been developed to interpolate point observations to a spatial framework (e.g. Thiessen 1911, Cressman 1959, Creutin and Obled 1982, Nuss and Titley 1994). Thus, comparing a simulated climatology to a climatology of observations interpolated to a grid has become common climate modeling practice (e.g. Labraga and López 1997, Bauer et al. 2002, Dickinson et al. 2006).

We use a Cressman (1959) interpolation to project the long-term soil-temperature observations onto the CCSM grid between 25°E-160°E and 38°N-72°N and determine grid-cell averages. Cressman-interpolation can yield erroneous values in mountainous terrain and may introduce redundancy if two sites are close together (e.g. Dingman 1994); however, in Russia, the number of sites in more complex terrain is low and sites are far from each other so redundancy plays no role. Simulated soil temperatures are

interpolated with a weighted distance to observational levels. Note that errors due to vertical interpolations are less than 0.1 K (e.g. Narapusetty and Mölders 2006).

Monthly grid-cell averages are calculated. The data permit us to determine three 30-year climatologies, 1951-1980, 1961-1990, and 1971-2000 (denoted first, second, and third climatology, hereafter) for evaluation of CCSM soil climatologic performance. Due to uncertainty introduced by observational procedure and interpolation, we determine good agreement between simulated and observed soil climatology if differences are ±1 K or less.

For 1951-2000, monthly precipitation data are available on a 2.5°x2.5° grid from the Global Precipitation Climatology Centre (GPCC). These data are based on quality-controlled and homogenized time-series from 9343 stations worldwide. After interpolating these GPCC data to the 2.8°x2.8° CCSM grid, three climatologies are calculated for Russia.

Snow-depth data are available for 1978-2003 from in-situ measurements at over 3000 stations. We project these data onto the CCSM grid using a Cressman (1959) interpolation. In addition to the aforementioned shortcomings of this interpolation, one has to be aware that mountain sites are not very representative for a large area (e.g. Frei and Schär 1998, Colle et al. 2000).

Cloud-cover-fraction data are available for 1984-2004 on a 2.5°x2.5° grid from the International Satellite Cloud Climatology project (ISCCP; Rossow and Schiffer 1999) and interpolated to the 2.8°x2.8° CCSM grid. ISCCP total cloud-cover fraction is the percentage of pixels identified as cloudy within a grid-cell.

Recently, various centers have worked on providing reanalysis data (e.g. National Centers for Environmental Prediction [NCEP], National Center for Atmospheric Research [NCAR], Eupropean Center for Medium Range Weather Forecast [ECMWF]). A reanalysis uses observational data from various sources, for instance, synoptic stations, ship observations, radiosondes, pibal, aircraft, radiance data from various satellite-borne radiometers (e.g. TOVS, TOMS, SSM/I, VTPR, geostationary satellites) and other data. All data used are quality assessed/quality controlled. A frozen state-of-the-art global

NWPM analysis/forecast system and a database as complete as possible are used to perform data assimilation to produce a record of global analyses of atmospheric fields (for details on reanalysis see e.g. Kalnay et al. 1996, Mitchell et al. 2004, Uppala et al. 2005). Thus, these reanalysis data can be considered as "observations" that are intelligently and physically and thermodynamically consistently "interpolated/balanced" to a horizontal and vertical grid by making use of the physical process relationships among the various quantities (e.g. Simmons et al. 2004, Uppala et al. 2005). Evaluating simulations using reanalysis data is common practice in modeling (e.g. Anthes 1983, Schmidt et al. 2006). We use ERA40-reanalysis 2.5°x2.5° data for evaluating the impact of inaccurately-simulated near-surface air temperatures on soil-temperature simulations because they have a higher temporal resolution than Climatic Research Unit (CRU; e.g. Mitchell et al. 2004) 0.5° data (Uppala et al. 2005, Drobot et al. 2006). ERA40 data and CRU data are highly correlated and do not significantly differ statistically (Drobot et al. 2006). Although ERA40 near-surface temperatures have positive bias in winter over the northern hemisphere in comparison with CRU data, the average bias is less than 3 K over Russia, and very small in summer (Hagemann et al. 2005). Correlation between ERA40 and NCEP/NCAR reanalysis is 0.998 globally and 0.996 for Russia; the bias between ERA40 and NCEP/NCAR reanalysis is -0.16 K for Russia, i.e. they provide similar results. Since the reanalysis data from different sources differ only marginally, we restrict our discussion to ERA40 data.

4.2.3 Simulations

CCSM is run in fully-coupled mode with 26 vertical layers at a spectral truncation of T42 corresponding to a spatial resolution of ≈2.8°x2.8°. CCSM is started with the ecliptic conditions of 1-1-1950 and CO_2 concentration of 355 ppmv. Each model component is spun up separately. Based on these simulations, we determine three climatologies, 1951-1980, 1961-1990, and 1971-2000.

Uncertainties in soil parameters (e.g. porosity, pore-size distribution index, saturated hydraulic conductivity) can be as large as the parameters themselves (e.g. Clapp and

Hornberger 1978, Cosby et al. 1984) and can cause statistical errors in simulated soil temperatures that can be especially great if soil water freezes or frozen ground thaws (Mölders et al. 2005). To assess the impact of soil-parameter inaccuracy on soil temperatures simulated by the fully-coupled CCSM, we perform sensitivity studies wherein we reduce (enhance) the sand fraction by 10 % and increase (decrease) the clay fraction by 10 % in Russia. These simulations are denoted S10D and S10I, respectively.

To investigate uncertainty resulting from differences between the vegetation in CCSM and in the grass-plot under which the measurements were made, we perform a simulation wherein forest is randomly replaced by grass and/or crops in some Russian grid-cells (total area change <1 %).

4.2.4 Analysis

Numerous studies have shown that CCSM's LSM reasonably well simulates land-surface processes, snow accumulation and snow-melt (Bonan et al. 2002), state variables, and fluxes in offline mode (Dai et al. 2003, Bonan and Levis 2006, Lannoy et al. 2006). To evaluate the coupled soil-temperature simulations we compare simulated and observed soil-temperature climatologies and identify reasons for discrepancies.

The main regulators of soil temperature are water-phase transitions, soil physical properties (primarily soil composition), rainfall and incoming energy and hence indirectly cloudiness, snow-depth, near-surface air temperature, and vegetation. Rainfall introduces heat into soil and refreshes total soil-water content. Snow cover insulates; too-thick or -thin simulated snow cover may yield an inaccurate upper boundary condition at the soil surface. The timing of snow-cover establishment in fall and disappearance in spring can similarly impact soil temperatures. Vegetation type and fraction may affect soil temperature by shading. Near-surface air temperatures influence soil temperatures via energy and water fluxes. Thus, several potential reasons for incorrectly simulated soil temperatures exist, namely incorrect atmospheric or soil forcing (near-surface air temperature, precipitation, snow depth), and differences between modeled and natural soil and/or vegetation type.

To attribute reasons for differences between simulated and observed soil temperatures, we compare near-surface air temperature climatologies calculated from CCSM data with those from ERA40 reanalysis data, CCSM-derived precipitation climatologies with those gained from GPCC data, CCSM-derived snow-depth climatologies with available snow-depth data, and CCSM-simulated cloud fraction climatologies with ISCCP data. Since snow-depth and cloud-fraction data are not available for a full 30-year period, and ERA40 data are not available for all three climatologies, we examine whether simulated data for the shorter period of data availability differ significantly from the three climatologies. The null hypothesis is that the "shorter CCSM dataset" and the CCSM climatologies do not differ at the 95 % or higher confidence level. This means we can use the available snow-depth data, ERA40 near-surface air temperature and ISCCP data to assess CCSM's performance in simulating these quantities for all three climatologies. The null hypothesis was confirmed in all cases.

Systematic and non-systematic errors can contribute to any simulation error. To evaluate performance and attribute error sources at different spatial and temporal scales, we calculate performance measures (e.g. Anthes 1983, Anthes et al. 1989) for the annual course and mean for all climatologies for both domain-averages and each grid-cell. The BIAS

$$\bar{\phi} = \frac{1}{n}\sum_{i=1}^{n}\phi_i \tag{4.1}$$

gives systematic errors from consistent misrepresentation of geometrical, physical, or numerical factors (e.g. difference in terrain elevation, soil or vegetation type, pipe heat conduction in the uppermost levels), while the standard deviation of error

$$\mathrm{SDE} = \left(\frac{1}{n-1}\sum_{i=1}^{n}\left(\phi_i - \bar{\phi}\right)^2\right)^{1/2} \tag{4.2}$$

represents random errors caused by uncertainty in initial and boundary conditions or observations (e.g. constant initial soil-temperature distribution worldwide, constant heat flux at the lower soil-model boundary, sampling density). Here ϕ_i denotes the difference

between simulated and "observed" soil temperature for the i^{th} grid-cell and n is the number of grid-cells over Russia. The root-mean-square error

$$\mathrm{RMSE} = \left(\frac{1}{n-1} \sum_{i=1}^{n} (\phi_i)^2 \right)^{1/2} \tag{4.3}$$

evaluates overall performance. Monthly mean values are used in the evaluation.

To test the hypothesis that incorrectly-simulated precipitation, snow depth, and/or cloud fraction cause errors in simulated soil temperatures we apply a Student's t-test to the differences between simulated and observed forcing quantities and the differences between simulated and observed soil temperatures. Note that in the following the word *significant* will only be used if differences or correlation pass a t-test at the 95 % or higher confidence level.

4.3 Results

4.3.1 Climatology 1951-1980

PILPS showed that most LSMs capture the phase of seasonal average soil temperature well for upper layers; most LSMs participating in PILPS 2d succeeded in simulating annual observed soil temperature variability in offline mode (Luo et al. 2003). As pointed out above, CCSM's LSM performs well in offline mode (Bonan et al. 2002, Dai et al. 2003, Bonan and Levis 2006, Lannoy et al. 2006).

The fully-coupled CCSM captures upper soil temperature phase well, and captures annually-averaged soil temperature reasonably well at all depths. However, CCSM fails to exactly capture amplitude (Fig. 4.1), providing soil temperatures that are slightly too high. At 0.2 m (0.4 m, 0.8 m, 1.6 m, 3.2 m) depth, the RMSE and greatest discrepancy amounts are 4.5 (4.0, 3.6, 3.7, 3.2) K and 8.1 (7.1, 5.9, 5.8, 4.9) K, respectively. Except at 3.2 m depth, the greatest errors occur at temperatures <275 K.

On the annual cycle, CCSM simulates thaw-up too early (Fig. 4.2). The simulated annual temperature wave is in phase with summer observations except at 1.6 m depth, where the minimum is slightly delayed by one month. Soil penetration of the simulated annual temperature curve is too flat for the cold portion of the amplitude curve and

slightly too deep for the warm portion. Consequently, CCSM overestimates 0.2 m soil temperature for most grid-cells from December to March (Fig. 4.2) and on average (Table 4.1). Performance at 0.4 and 1.6 m depth is similar, but frequency decreases marginally with increasing depth. At 0.8 and 3.2 m, CCSM overestimates soil temperature year-round (Fig. 4.2). At all depths, discrepancies increase with decreasing temperature (Figs. 4.1, 4.2). In April, at all depths, CCSM will overestimate soil temperature for most grid-cells if soil temperature is below freezing point, and underestimates it otherwise, leading to overall overestimation. In May and June, the general pattern shifts towards underestimation at 0.2 m, 0.4 m and 1.6 m depth (Fig. 4.2). The best agreement between simulated and observed climatology occurs around the freezing point. In July and August, at 0.8 m depth and above, CCSM tends to overestimate soil temperatures by <5 K at the warmer ($T > 279\,K$) end of the temperature range, and underestimate by <7 K at the cooler end ($T \leq 279\,K$). On average, CCSM underestimates soil temperatures by up to 1.2 K at 0.2 m and overestimates them by up to 0.5 K at 0.4 and 0.8 m depth in these months (Fig. 4.2). At 1.6 m, the tendency to overestimate soil temperature is obvious for the colder (<275 K), but less obvious for the warmer (>285 K) end of the range. In September, CCSM underestimates soil temperatures at 0.2 and 0.4 m depth by up to 0.6 K and 0.4 K for most grid-cells. At 0.8 m and 1.6 m, CCSM overestimates soil temperatures by up to 0.4 K and 1.1 K, respectively, for most grid-cells (Fig. 4.2). In October, the general pattern again shifts towards overestimation by up to 2.5 K. In November, simulated soil conditions are too warm for soil temperatures below the freezing point at all depths (Fig. 4.2).

The RMSEs decrease with increasing depth and are greater in winter than in the other seasons (Fig. 4.3). In winter (summer), they reach up to 8.1 K (3.2 K) in the uppermost layer. Obviously secondary maxima of high RMSEs exist around 1.6 m depth from April to May, and August to December. The lowest RMSEs (1.9 K) exist in the upper soil in October (Fig. 4.3).

The overall soil-temperature overestimation results in overall positive BIAS (Fig. 4.4) indicating overall systematic error. At 0.2 m depth, BIAS is positive in fall and spring, but negative in summer. Uppermost layer soil temperatures exhibit maximum positive and negative BIAS in January (7.3 K) and August (-1.2 K). In these months, systematic error from heat conduction by the ebonite pipe is largest because air temperatures are at their annual extremes. BIAS decreases with increasing depths (Fig. 4.4) because most variability and conduction occurs close to the surface. At 0.8 m depth, soil temperatures are positively biased year-round with smaller BIAS in summer. Systematic error may also arise because single-point soil-temperature measurements were taken under snow-covered grass-plots, while model soil temperatures are simulated under various PFTs and weighted according to grid-cell PFT-fraction to obtain the grid-cell soil temperature. Differences between observation-site elevation and grid-cell-averaged terrain height mean a consistent misinterpretation of geographical factors with the most severe consequences in transition seasons, when elevation may determine whether snow remains (spring), or has begun to fall (autumn). Note that first snow may fall as early as August in the northern and mountainous parts of the domain.

At 0.2 m, average SDEs are highest in winter (up to 3.1 K in January) and lowest (1.6 K) in October. The 0.4 m depth is similar, with marginally smaller SDEs than at 0.2 m depth. At 0.8 m, the maximum (minimum) SDE occurs in July (November); the maximum (minimum) 1.6 m SDE occurs in August (November). These results indicate that random errors vary with season. As shown by Mölders and Romanovsky (2006), simulated soil-temperature accuracy depends, among other things, on the soil model's lower boundary depth and the boundary condition chosen. Thus, the constant-heat-flux lower-boundary condition used in CCSM may be more representative of actual conditions in summer than in winter.

At all depths, simulated and observed climatology correlate most strongly in summer. The highest correlation (0.952) occurs in September in the uppermost layer; the lowest correlation (0.734) exists in January at 0.4 m depth.

4.3.2 Climatology 1961-1990 and 1971-2000

In the second and third climatologies, mean annual soil temperatures are overestimated by 2.5 K, on average (Table 4.1). Similar to the first climatology, soil temperatures are overestimated from September to June, and simulated best for July and August. Simulated soil temperatures are slightly closer to observations in the third, than in the first or second climatology, probably due to the change in observational time and frequency that occurred in 1965.

For the second (third) climatology, at 0.2, 0.4, 0.8, 1.6, 3.2 m depth, RMSEs amount 4.4, 3.9, 3.5, 3.6, 3.1 K (4.4, 3.9, 3.5, 3.6, 3.2 K) and the largest discrepancy amounts 7.8, 6.8, 5.6, 5.4, 4.8 K (8.0, 6.9, 5.8, 5.6, 4.9 K). For all three climatologies, the highest RMSEs, SDEs, and BIASes occur in the uppermost 0.2 m because temporal variability is greater close to the surface than deeper in the soil (Figs. 4.3, 4.4). All three climatologies show positive BIAS from October to April and negative BIAS from May to September (Fig. 4.4), i.e. systematic errors are the same for all climatologies.

In summary, for all three climatologies CCSM captures the annual average soil temperature reasonably well, but simulates summer soil temperature better than winter soil temperature. Because of the acceptable RMSEs and high correlation, we conclude that the fully coupled CCSM acceptably simulates soil temperatures (Figs. 4.3, 4.4).

4.4. Discussion

Errors in precipitation and/or near-surface air temperatures may cause underestimates in summer; discrepancies between simulated and actual snow-cover conditions may cause overestimates in winter. Inaccurate cloud-fraction simulation may influence surface and soil temperature and contribute to discrepancies between simulated and observed soil temperatures. In snow-free months, differences in surface conditions for measurements (grass-plot) and model (soil temperature grid-cell-average weighted for the various surface tiles within the grid-cell) may cause discrepancies. Inaccurate soil physical parameters may play a role. In the following sections, we attribute the reasons for errors in simulated soil temperature.

4.4.1 Near-surface air temperatures

To test the hypothesis that incorrectly simulated near-surface air temperatures contribute to simulated soil-temperature errors, we compare CCSM simulated near-surface air temperature with ERA40 near-surface air temperature. Since the ERA40 data start in 1958, only the second and third climatology can be used for this purpose. As pointed out above, the climatologies do not differ significantly from each other; therefore results for the second and third climatologies can be generalized for the first climatology.

Overall, CCSM simulates annually-averaged near-surface temperature well with an overall RMSE of 3.0 K (3.0 K), BIAS of -1.1 K (-1.0 K), and SDE of 2.6 K (2.5 K) for the second (third) climatology (Table 4.1). For both near-surface air temperature climatologies, winter RMSEs and SDEs exceed those obtained for summer (Fig. 4.5). Analysis of RMSEs and correlations between near-surface temperature climatologies derived from CCSM and ERA-40 also indicate better summer than winter performance. For some grid-cells, CCSM slightly overestimates near-surface air temperatures in winter yielding positive BIAS, but underestimates them for other seasons resulting in negative BIAS (Fig. 4.5). In summer, negative BIAS increases due to systematic errors caused by misrepresentation of convective events (see section 4.4.2). Note that, in CCSM, onset of daytime moist convection occurs about 4 hours too early and the diurnal cycle of convection is too smooth (Dai and Trenberth 2004).

To test the hypothesis that incorrect near-surface air temperatures contribute to soil temperature discrepancies, differences between CCSM and ERA40 near-surface air temperatures are compared to differences between simulated and observed soil temperature at the various soil depths (Table 4.2). These differences correlate positively and significantly except for the 1.6 m and 3.2 m depths in April and May. The correlation decreases with increasing depth in summer except at 3.2 m because near-surface air temperature has less impact on deeper than on upper soil layers. At 0.8-3.2 m depth, soil temperatures lag air temperatures by roughly 1-3 months in accord with observations (see e.g. Geiger 1961). Obviously, the lag is not well captured for the 3.2 m depth because of the lower constant flux boundary condition. Generally, correlation is lower in winter than

summer due to the insulating snow-pack effect on the upper soil, but it remains significant in winter.

In summary, CCSM captures near-surface temperatures reasonably well; correlation of near-surface air temperature and soil temperature errors is less in winter than summer. Thus, inaccuracies in simulated near-surface air temperatures contribute slightly to discrepancies between simulated and observed soil temperatures, but are not the major contributor, especially in winter.

4.4.2 Cloud fraction

To test the hypothesis that incorrect cloud fraction contributes to soil-temperature discrepancies we compare ISCCP and CCSM cloud-fraction climatologies. CCSM overestimates cloud fraction from October to March and underestimates it otherwise. Overestimation of cloud fraction increases net radiation and downward long-wave radiation, warming the near-surface air (Fig. 4.6). This agrees broadly with findings by Walsh and Chapman (1998), who compared observational with atmospheric reanalysis data. In summer, convective clouds occur more often. The simulated diurnal convective-cloud cycle is too smooth (Dai and Trenberth 2006), possibly overestimating irradiation later in the day. Thus, underestimated cloud fraction may contribute to summer soil-temperature underestimates. However, another effect of overestimated cloud cover is the reduction of incoming solar radiation that may reduce surface heating. The resulting cooling of the surface may reduce outgoing radiation. Which of the two effects prevails depends on the region and season.

Over Russia, CCSM overestimates cloud fraction by up to 12 % from October to April and underestimates it by up to 11.5 % at other times (Fig. 4.6). RMSEs between simulated and observed cloud fraction are comparatively lower in spring and fall than in other seasons. Correlations are highest in summer (0.871). Nevertheless, over southern Russia simulated and observed cloud fraction differ most in July (Fig. 4.6).

Cloud-fraction discrepancies correlate positively with soil-temperature discrepancies from October to April, and negatively for the other months (Fig 4.6). Correlations are

significant except for October. Similar correlation behavior (positive from November to April, negative otherwise) with significant correlation year-round was found between discrepancies in near-surface air temperature and in cloud fraction. This means that incorrectly-simulated cloud fraction may yield incorrect near-surface temperatures and soil temperatures. Negative correlation magnitudes exceed positive correlation magnitudes. Cloud-fraction discrepancies correlate positively with precipitation discrepancies. Correlations are significant except for February to April, and are highest in late summer/early fall. These findings suggest that convective parameterization shortcomings may contribute to incorrectly-simulated summer soil temperatures.

4.4.3 Precipitation

CCSM provides realistic distributions for precipitation >1 mm, but simulates rain too frequently and at reduced intensity in summer (Dai and Trenberth 2004). Observed annual precipitation amounts over Russia from GPCC data are 441.8, 438.8, and 436.8 mm/y for the first, second, and third climatology, while CCSM provides 524.4, 523.8, and 484.0 mm/y, overestimating precipitation by similar amounts for all climatologies. From October to May CCSM overestimates monthly precipitation by up to 18 mm/mon; from June to September it underestimates precipitation by up to 12 mm/mon for all three climatologies. Nevertheless, CCSM captures minimum and maximum precipitation well (Fig. 4.7).

CCSM underestimates annual accumulated precipitation by up to 10 mm/mon in southwestern (40°N-45°N, 35°E-50°E) and southeastern (38°N-42°N, 120°E-132°E) Russia (Fig. 4.8). CCSM captures annual accumulated precipitation well in central Russia (55°N-65°N, 55°E-95°E) but overestimates this quantity elsewhere. Discrepancies between simulated and observed precipitation climatology mainly occur in mountainous and lake-rich regions. The reasons are manifold. CCSM treats lakes as a percentage area within a grid-cell no matter of their real extent and where they are located therein. It also ignores mesoscale circulations that may establish in response to the lake-land temperature and moisture differences. Therefore, the location of water supply to the atmosphere

differs between the model and nature with consequences for cloud and precipitation formation (for a detailed discussion of the consequences of this mosaic-type approach see Mölders et al. (1996) or Giorgi and Avissar (1997)). In CCSM terrain elevation is grid-cell average height, so mountains are flatter than the highest natural peaks. Consequently, orographically-induced precipitation may be underestimated or occur further downwind than in nature. Furthermore, sites in mountainous terrain are representative of a smaller area than sites in flat homogeneous terrain (e.g. Frei and Schär 1998, Colle et al. 2000).

Precipitation underestimation yields negative BIAS from June to September for all climatologies, and positive BIAS for the other months (Fig. 4.7). The maximum positive and negative Russian domain-average BIAS is about 18 mm/mon in winter and -12 mm/mon in summer, respectively. During winter BIAS increases for all climatologies indicating an increase in systematic errors. Systematic errors result from, among other things, misrepresentation of the partitioning of solid and liquid precipitation in the cloud parameterization, assumptions about snow-crystal shape and size, vertical grid resolution, and incorrectly-simulated atmospheric moisture transport. Moisture transport can be wrong due to inaccurately-simulated atmospheric moisture content (water vapor, cloud and precipitation particles), wind speed and direction. Note that since cloud- and precipitation-formation are subgrid-scale processes for any GCM they must be parameterized; i.e. these errors cannot be avoided, but must be minimized. Potential consequences of CCSM's cloud parameterizations for simulated precipitation have been discussed above.

SDEs are higher in summer than in winter (Fig. 4.7). SDEs increase as summer progresses, because convective precipitation becomes more likely. Since convective precipitation is very local, errors in grid-cell-averaged observed precipitation may result from the precipitation network not capturing a precipitation event, and from the difficulty of interpolating convective precipitation onto a grid (e.g. Dingman 1994). Winter precipitation is spatially more homogeneous than summer precipitation. Thus, SDEs mainly result from catch deficiencies that are greater for solid than for liquid precipitation. Depending on wind speed, snowfall-catch deficiencies can cause measured

precipitation errors that may exceed 30 % of the snowfall (e.g. Dingman 1994, Yang and Woo 1999, Yang et al. 2000, Sugiura et al. 2003).

SDEs are higher (<15 mm/month) in southwestern (40°N-58°N, 30°E-90°E) and eastern (40°N-50°N, 110°E-140°E) Russia (Fig. 4.8). High SDEs mainly occur in mountainous, coastal, and lake-rich regions due to coarse network density in these areas.

Analysis of RMSEs and correlations between simulated and observed precipitation also manifests better winter than summer model performance for all three climatologies (Fig. 4.7). RMSEs amount to 20 mm/mon, on annual average, but reach up to 25.9 mm/mon in summer.

To test the hypothesis that incorrectly-simulated precipitation causes errors in simulated soil temperatures, differences between simulated and observed precipitation are compared to differences between simulated and observed soil-temperature (Fig. 4.9). Since the active layer is typically saturated except close to the surface (Hinkel et al. 2001), we only consider precipitation-error impacts on soil-temperature errors for the upper 0.2 m of soil, assuming that heat conduction to deeper layers by precipitation is small. Typically this layer is non-frozen in summer.

Differences are negatively correlated for summer (Table 4.3). Though low, this correlation is significant. Therefore if CCSM overestimates precipitation, it will underestimate soil temperatures and vice versa (e.g. Figs. 4.4, 4.6, 4.8). In winter, incorrectly-simulated precipitation (precipitation occurring in a given winter month, not snow already accumulated on the ground) and errors in simulated soil temperature are uncorrelated, because the snow-pack is already thick enough to be an effective insulator. Errors in simulated precipitation that occur in fall affect accumulated snow depth and hence soil temperature (see section 4.4.4). The influence of errors in simulated near-surface air temperature (and hence loss of snow depth due to sublimation) have been discussed in section 4.4.1.

Our findings identify inaccurately-simulated precipitation as contributing to inaccurate soil-temperature simulations in summer. As expected, errors in precipitation

simulation can be excluded as a major reason for inaccurate winter soil temperatures once the snow-pack exceeds critical thickness.

4.4.4 Snow depth

The analysis shows little discrepancy between simulated and observed snow depth for November to March. Simulated 1979-2003 snow-depth climatology strongly correlates with the three climatologies (r_1=0.964, r_2=0.983, r_3=0.995, where r_n and n are the correlation coefficient and n^{th} climatology, respectively). As mentioned, simulated 1979-2003 climatology does not differ significantly from the three climatologies; therefore we can use the 1979-2003 snow-depth observations to (1) assess CCSM's snow-depth simulation performance and (2) test the hypothesis that simulated snow-depth errors may be a major cause for inaccurately-simulated soil temperatures.

On average, CCSM overestimates snow depth, but captures the temporal evolution of the snow-pack well (Fig. 4.10). According to a Student's t-test and F-test CCSM fails to capture snow-depth distribution and variability in the same locations where it failed for precipitation (therefore not shown), for the same reasons. Therefore, we exclude mountainous and coastal regions from further investigation; in summer, the remaining Russian areas are snow-free.

CCSM's overestimation of near-surface winter air temperatures increases potential evaporation. Consequently, more snow will sublimate if air is sub-saturated with respect to ice. Enhanced water-vapor supply affects cloud and precipitation formation. Snow depth can also be affected if water vapor entering the atmosphere falls in the same region as snow. Since snow depth is an accumulated quantity, errors (RMSE, BIAS, SDE) increase from November to April; systematic errors from differences between model and real terrain height and lake treatment also accumulate over winter. BIASes and SDEs are largest in southwestern (50°N-70°N, 35°E-90°E), RMSEs in western and southwestern (50°N-70°N, 35°E-90°E) Russia (Fig. 4.11). Catch deficits, relatively low sampling density, and errors in simulated near-surface air temperatures and snowfall contribute to SDEs.

Differences between simulated and observed snow depth are compared with differences between simulated and observed soil temperature at all soil depths to test the hypothesis that incorrect snow depth causes errors in simulated soil-temperature climatology (Fig. 4.12). These differences correlate positively in winter (Table 4.4), i.e. if CCSM overestimates (underestimates) snow depth, it will overestimate (underestimate) soil temperature. Correlations, while low, are significant in winter. Correlations decrease from November to March at all depths because once the snow-pack reaches a certain thickness, its insulating effect hardly differs for increases in snow depth, and soil temperatures are decoupled from atmospheric temperatures. Generally, snow-depth effects become less pronounced with snow depths >0.4 m (see Zhang et al. 2005). This critical snow depth may vary spatially.

In late fall, however, a too-thin or too-thick snow-pack affects simulated soil temperature. Since soil has a "memory", errors in soil temperatures are still visible in winter. Note that soil temperatures simulated without and with consideration of a 2-day snow-cover of a few centimeters differ at 1m depth long after the snow event (Mölders et al. 2003); delaying snow-cover onset by 10 days can decrease maximum and annual mean soil temperatures by up to 9.0 K and 0.7 K at the surface, and 1.1 K and 0.4 K at 2.0 m depth (Ling and Zhang 2003). We conclude that incorrect simulation of autumn snowfall amount and snow depth explains some of the errors in simulated winter soil-temperature climatology. A sensitivity study using satellite-derived snow-depth data provides results and conclusions similar to those gained from ground-based measurements (therefore not discussed).

4.4.5 Sensitivity studies

We test the hypothesis that incorrect soil parameters explain differences between simulated and observed soil temperatures using results of the S10D, S10I, and original simulation. Averaged over Russia, the altered sand percentage increases (decreases) porosity from 0.434 to 0.439 (0.429 m^3/m^3) and increases (decreases) saturated soil heat capacity about 1 % in S10D (S10I). Increased heat capacity implies that more energy is

needed to increase soil temperature. Decreased soil thermal conductivity results in reduced upward soil-heat flux in winter; in summer, decreased thermal conductivity reduces downward soil-heat flux, i.e. soil heating. However, summer and winter effects are not equal because summer conditions usually lasts 3-4 months, compared to 8-9 months of winter. Consequently, on average, less energy leaves the soil leading to higher soil temperatures for S10D than the reference run. The same argument, with opposite sign, applies for S10I where soil temperature decreases; i.e. in S10D (S10I), soil temperatures are higher (lower) than temperatures obtained using original soil properties. Altered soil temperatures result in corresponding slightly-altered temporal and spatial distributions of RMSE, BIAS, SDE, and correlation coefficients. Soil temperatures obtained from S10D, however, differ significantly from reference-simulation temperatures in December at 0.2 m depth and from January to March at 0.2 m, 0.4 m, and 0.8 m. Soil temperatures obtained from S10I differ significantly from the reference simulation in July and August at all depths. Soil temperatures obtained from S10I differ significantly from those obtained by S10D in December and January at 0.2 m and 0.4 m depth, February and March at all depths, and in June, July, and August at 0.2 m, 0.4 m, and 0.8 m. Temporally and spatially each parameter set provides better results for one climatology. No obvious overall advantage for one or the other choice of soil parameters exists (therefore not shown). The same is true for other quantities (near-surface air temperatures, cloud fraction, precipitation, snow depth).

These findings indicate that soil characteristics contribute notably or even significantly to errors in simulated soil-temperature climatologies. The fact that decreasing or increasing sand percentage yields no clear result further suggests that better or worse results may occur by accident. Russian soils have high peat content, while CCSM assumes mineral soils. Obviously, one can better capture organic soil behavior in one or another season by assuming another mineral-soil type, but not over the annual cycle. Thus, including organic material is an urgent need for LSMs.

A sensitivity study performed with small vegetation disturbance to reflect differences between real-world and modeled vegetation shows that these differences marginally contribute to errors in simulated soil-temperature climatologies for the coupled model.

4.5 Conclusions

Soil temperatures simulated by the fully coupled Community Climate System Model version 3.0 are evaluated using 50 years of Russian soil-temperature measurements. This dataset permits construction of three soil climatologies (1951-1980, 1961-1990, 1971-2000). Observed snow-depth, cloud and precipitation climatology, ERA40 near-surface air temperature data and results from sensitivity studies performed with CCSM are used to explain and identify errors resulting from the soil model itself, and from running the soil model in fully-coupled mode within CCSM.

The fully-coupled CCSM captures the phase of soil-temperature wave well in upper soil, and the annual average soil-temperatures reasonably well at all depths; it fails to capture the amplitude exactly (Fig. 4.1). Generally, CCSM overestimates (underestimates) soil temperatures at all depths in winter (summer) with a better performance in summer than winter. In spring and fall simulated and observed climatologies agree the best.

BIASes and SDEs are higher in mountainous and lake-rich areas indicating that some discrepancies result from systematic and random errors caused by consistent model misinterpretation of physical, geometrical, and numerical aspects, and from the coarser network in these areas. Discrepancies between model and real terrain elevation and treatment of subgrid-scale heterogeneity, cloud and precipitation parameterizations, and pipe heat conduction cause these systematic errors. The first three reasons clearly relate to the coupling. Offline simulations use correct site elevation and observed forcing data. Systematic errors from pipe conduction affect offline and online evaluations. Upper-level soil temperature measurements are made using pipes exposed to the air; pipe heat conduction explains high BIAS in January and August when near-surface air temperatures reach their extremes. Other occasionally-significant (at the $\geq$95 % confi-

dence level) systematic error results from using mineral soil instead of organic soil physical parameters. This error is also coupling-independent because the soil model cannot consider organic components at all. Nevertheless, soil parameters are a challenge in any coupling since no "area-representative" soil physical parameters exist. Small differences between model and real vegetation marginally affect simulated soil temperatures, i.e. vegetation differences that play a great role in offline evaluations become relatively unimportant when comparing soil-temperature observations interpolated to a 2.8^{o}x2.8^{o}-grid with soil temperatures simulated by a fully-coupled model. The temporal behavior of SDEs suggests that CCSM's constant-heat-flux lower boundary condition represents actual conditions better in summer than winter, i.e. there seems to be a gradient in the winter flux. Therefore future development should address lower boundary condition formulation.

Comparing simulated near-surface air-temperature climatology and ERA40 near-surface air-temperature climatology shows that although CCSM tends to overestimate winter and underestimate summer near-surface air temperature, overall CCSM simulates near-surface air temperature reasonably. Errors in near-surface air temperatures and soil temperatures significantly correlate, but less strongly in winter than summer (Table 4.2), indicating that incorrectly-simulated atmospheric forcing contributes to, but is not the main reason for discrepancy between simulated and observed soil-temperature climatology.

Precipitation evaluation shows that CCSM overestimates precipitation from October to May and underestimates it for the other months. These discrepancies mostly occur in southeastern and southwestern Russia. The high RMSEs and SDEs of summer precipitation imply that CCSM fails to simulate summer precipitation adequately. Differences between simulated and observed precipitation significantly correlate with differences between simulated and observed soil temperature at 0.2 m depth in summer, i.e. inaccurate precipitation simulations significantly contribute to underestimating summer soil temperature.

Cloud fraction discrepancies correlate significantly with soil, near-surface air temperature, and precipitation discrepancies; the correlation is especially high in summer. Based on these correlations we conclude that convective parameterization shortcomings may be the main reason for incorrectly-simulated summer soil temperatures. Some soil-temperature error, therefore, results from running the soil model in a fully-coupled mode with the climate system model. Consequently, improving convective clouds and precipitation parameterizations is a pre-requisite for improved summer soil-temperature simulation by the fully-coupled CCSM.

CCSM overestimates winter precipitation and consequently snow depth according to satellite data and ground-based observations. Snow-depth RMSEs, SDEs, and BIASes increase in winter because snow depth is an accumulative quantity sensitive to a variety of processes. Incorrect simulation of fall snowfall amount and snow depth explains some errors in simulated winter soil-temperature climatology. Correlations between errors in simulated snow depth and simulated soil temperature are low, but significant in winter. Consequently, inaccurate simulation of snowfall and sublimation, and hence inaccurate snow depth are a cause of overestimating winter soil temperature. Surface heterogeneity parameterization and discrepancies between model and real world terrain heights, as well as lake treatment, are identified as sources for incorrect snow depth especially in fall; i.e. these errors relate to the coupled mode.

For now we conclude that soil temperatures simulated by the fully-coupled CCSM and, hence, any assessments about permafrost thawing are more uncertain in mountainous and lake-rich terrain, and where summer convection plays a dominant role. Thus, improving cloud and surface heterogeneity parameterizations and using a fine model resolution that better captures terrain height and coastlines are urgent needs for improving soil temperature simulations by the fully coupled CCSM.

Acknowledgements

We thank D. Atkinson, U.S. Bhatt, G. Kramm and the anonymous reviewers for fruitful comments, C. O'Connor for editing, C. Swingly for help with ERA40 data, and

B. Rudolf for access to Global Precipitation Climatology Center data. This research was supported by the International Arctic Research Center, University of Alaska Fairbanks, under the auspices of the NSF cooperative agreement OPP-0327664. Zhang, Etringer, and McCreight were partly funded through NSF grants OPP-0229766 and OPP-0353910. The National Center for Atmospheric Research and Arctic Region Supercomputing Center provided computational support. Soil temperature data are available from Zhang T, Barry RG, Gilichinsky D (2001) Russian historical soil temperature data, Boulder, Colorado, USA: National Snow and Ice Data Center, Digital media (http://nsidc.org/arcss078.html/).

References

Anthes RA (1983) Regional models of the atmosphere in middle latitudes. Mon Wea Rev 111: 1306-1335

Anthes RA, Kuo YH, Hsie EY, Low-Nam S, Bettge TW (1989) Estimation of skill and uncertainty in regional numerical models. Quart J Roy Meteorol Soc 111: 763-806

Bauer M, Genio AD, Lanzante JR (2002) Observed and simulated temperature-humidity relationship: sensitivity to sampling and analysis. J Climate 15: 203-215

Blackmon M, Boville B, Bryan F, Dickinson R, Gent P, Kiehl J, Moritz R, Randall D, Shukla J, Solomon S, Bonan G, Doney S, Fung I, Hack J, Hunke E, Hurrell J, Kutzbach J, Meehl J, Otto-Bliesner B, Saravanan R, Schneider EK, Sloan L, Spall M, Taylor K, Tribbia J, Washington W (2001) The Community Climate System Model. Bull Amer Meteorol Soc 82: 2357-2376

Bonan GB (1996) The NCAR Land Surface Model (LSM version 1.0) coupled to the NCAR Community Climate Model. NCAR Tech. Note NCAR/TN-429+STR 171pp

Bonan GB, Oleson KW, Vertenstein M, Levis S, Zeng X, Dai Y, Dickinson RE, Yang Z-L (2002) The Land Surface Climatology of the Community Land Model Coupled to the NCAR Community Climate Model. J Climate 15: 3123-3149

Bonan GB, Levis S (2006) Evaluating aspects of the Community Land and Atmosphere Models (CLM3 and CAM3) using the CLM's dynamic global vegetation model. J Climate 19: 2290-2301

Boone A, Masson V, Meyers T, Noilhan J (2000) The influence of the inclusion of soil freezing on simulations by a soil-vegetation-atmosphere transfer scheme. J Appl Meteor 39: 1544-1569

Bowling LC, Lettenmaier DP, Nijssen B, Graham LP, Clark DB, Maayar M, Essery R, Goers S, Gusev YM, Habets F, Hurk B, Jin J, Kahan D, Lohmann D, Ma X, Mahanama S, Mocko D, Nasonova O, Niu GY, Samuelsson P, Shmakin AB, Takata K, Verseghy D, Viterbo P, Xia Y, Xue Y, Yang ZL (2003) Simulation of high latitude hydrological processes in the Torne-Kalix basin: PILPS Phase 2(e) 1:

Experiment description and summary intercomparisons. Global and Planetary Change 38: 1-30

Briegleb BP, Bitz CM, Hunke EC, Lipscomb WH, Holland MM, Schramm JL, Moritz RE (2004) Scientific description of the sea ice component in the Community Climate System Model, version three. NCAR Technical Note NCAR/TN463-STR, pp 70

Chen F, Dudhia J (2001) Coupling an advanced land surface-hydrology model with the Penn State-NCAR Modeling System. Part I: Model implementation and sensitivity. Mon Wea Rev 129: 569-585

Cherkauer KA, Lettenmaier DP (1999) Hydrologic effects of frozen soils in the upper Mississippi River basin. J Geophys Res 104D: 19599-19610

Clapp RB, Hornberger GM (1978) Empirical equations for some soil hydraulic properties. Water Resour Res 14: 601-604

Colle BA, Mass CF, Westrick KJ (2000) MM5 precipitation verification over the Pacific Northwest during the 1997-99 cool seasons. Wea Forecasting 15: 730-744

Collins WD, Rasch PJ, Boville BA, Hack JJ, McCaa JR, Williamson DL, Kiehl JT, Briegleb BP, Bitz C, Lin SJ, Zhang M, Dia Y (2004) Description of the NCAR Community Atmospheric Model (CAM3). Technical report NCAR/TN464+STR, National Center for Atmospheric Research, Boulder, CO, pp 226

Collins WD, Bitz CM, Blackmon ML, Bonan GB, Bretherton CS, Carton JA, Chang P, Doney SC, Hack JJ, Henderson TB, Kiehl JT, Large WG, McKenna DS, Santer BD, Smith RD (2006a) The Community Climate System Model: CCSM3. J Climate 19: 2122-2143

Collins WD, Rasch PJ, Boville BA, Hack JJ, McCaa JR, Williamson DL, Briegleb BP, Bitz CM, Lin SJ, Zhang M (2006b) The formulation and atmospheric simulation of the Community Atmosphere Model: CAM3. J Climate 19: 2144-2161

Cosby BJ, Hornberger GM, Clapp RB, Ginn TR (1984) A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils. Water Resour Res 20: 682-690

Cressman GP (1959) An operational objective analysis system. Mon Wea Rev 87: 367-374

Creutin JD, Obled C (1982) Objective analyses and mapping techniques for rainfall fields: an objective comparison. Water Resources 18: 413-431

Dai A, Trenberth KE (2004) The diurnal cycle and its depiction in the Community Climate System Model. J Climate 17: 930-951

Dai Y, Zeng X, Dickinson RE, Baker I, Bonan G, Bosilovich M, Denning S, Dirmeyer P, Houser P, Niu G, Oleson K, Schlosser A, Yang ZL (2003) The Common Land Model (CLM). Bull Amer Meteor Soc 84: 1013-1023

Dickinson RE, Oleson KW, Bonan G, Hoffman F, Thornton P, Vertenstein M, Yang Z-L, Zeng X (2006) The Community Land Model and its climate statistics as a component of the Community Climate System Model. J Climate 19: 2302-2324

Dingman LS (1994) Physical Hydrology, Prentice Hall, New York, pp. 575

Drobot S, Maslanik J, Herzfeld UC, Fowler C (2006) Uncertainty in temperature and precipitation datasets over terrestrial regions of western Arctic. Earth Interaction 10: 1-17

Esch DC, Osterkamp TE (1990) Cold regions engineering: climatic warming concerns for Alaska. J Cold Regions Engineering 4: 6-14

Frei C, Schär C (1998) A precipitation climatology of the Alps from high-resolution rain-gauge observations. Int J Climatol 18: 873-900

Friedrich K, Mölders N (2000). On the influence of surface heterogeneity on latent heat-fluxes and stratus properties. Atmos Res 54: 59-85

Fuchs M, Campbell GS, Papendick RI (1978) An analysis of sensible and latent heat flow in a partially frozen unsaturated soil. Soil Sci Soc America J 42: 379-385

Geiger R (1961) Das Klima der bodennahen Luftschicht, Vieweg, Braunschweig, pp 646

Giorgi F, Avissar R (1997) Representation of heterogeneity effects in earth system modeling: Experience from land surface modeling. Rev Geophys 35: 413-438

Gilichinsky DA, Barry RG, Bykhovets SS, Sorokovikov VA, Zhang T, Zudin SL, Fedorov-Davydov DG (1998) A century of temperature observations of soil climate: methods of analysis and long-term trends. In: Lewkowicz AG, Allard M (eds) The Proceedings of the 7th International Conference on Permafrost, Yellowknife, Canada, June 23-27, 1998. 313-317

González-Rouco F, von Storch H, Zorita E (2003) Deep soil temperature as proxy for surface air-temperature in a coupled model simulation of the last thousand years. Geophys. Res. Lett 30: 2116, doi:10.1029/2003GL018264

Grayson RB, Western AW (1998) Towards areal estimation of soil water content from point measurements: time and space stability of mean response. J Hydrol 207: 68-82

Hack J (1994) Parameterization of moist convection in the National Center for Atmospheric Research community climate model (CCM2). J Geophys Res 99: 5551-5568

Hagemann S, Arpe K, Bengtsson L (2005) Validation of the hydrological cycle of ERA40, ECMWF ERA-40 Project Report. Series 24, Reading, UK, pp42

Henderson-Sellers A, Yang ZL, Dickinson RE (1993) The Project for Intercomparison of Land-surface Parameterization Schemes. Bull. Am. Meteorol. Soc. 74: 1335-1349

Henderson-Sellers A, Pitman AJ, Love PK, Irannedjad P, Chen TH (1995) The Project for Intercomparison of Land surface Parameterization Schemes (PILPS): Phase 2 and 3. Bull Amer Meteorol Soc 76: 489-503

Hinkel KM, Paetzold RF, Nelson FE, Bockheim JG (2001) Patterns of soil temperature and moisture in the active layer and upper permafrost at Barrow, Alaska: 1993-1999. Global and Planetary Change 29: 293-309

Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J, Zhu Y, Leetmaa A, Reynolds B, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo KC, Ropelewski C, Wang J, Jenne R, Josep D (1996) The NCEP/NCAR 40-year reanalysis project. Bull Amer Meteor Soc, 77: 437-470

Kane DL, Hinzman LD, Benson CS, Liston GE (1991) Snow hydrology of a headwater Arctic Basin, Physical measurements and process studies. Water Resour 27: 1099 1109, 10.1029/91 WR00262

Kiehl JT, Shields CA, Hack JJ, Collins W (2006) The climate sensitivity of the Community Climate System Model version 3 (CCSM3). J Climate 19: 2584-2596

Koren V, Schaake J, Mitchell K, Duan Q-Y, Chen F, Baker JM (1999) A parameterization of snowpack and frozen ground intended for NCEP weather and climate models. J Geophys Res 104: 19,569-19,585

Labraga JC, López M (1997) A comparison of the climate response to increased carbon dioxide simulated by the General Circulation Models with mixed layer and dynamics ocean representation in the region of South America. Int J Climatol 17: 1635-1650

Lannoy GD, Houser PR, Pauwels VRN (2006) Assessment of model uncertainty for soil moisture through ensemble verification. J Geophys Res 111, D10101, doi:10.1029/2005JD006367

Ling F, Zhang T (2003) Impact of the timing and duration of seasonal snow cover on the active layer and permafrost in the Alaskan Arctic. Permafrost Periglacial Processes 14: 141-150

Luo L, Robock A, Vinnikov KY, Schlosser CA, Slater AG, Boone A, Braden H, Cox P, deRosany P, Dickinson E, Dai Y, Diuan Q, Etchevers P, Henderson-Sellers A, Gedney N, Gusev YM, Habets F, Kim J, Kowalczyk E, Mitchell K, Nasonova ON, Noilhan J, Pitman AJ, Schaake J, Shmakin AB, Smirnova TG, Wetzel P, Xue Y, Yang Z-L, Zeng QC (2003) Effects of frozen soil on soil temperature, spring infiltration, and runoff: Results from the PILPS2(d) experiment at Valdai, Russia. J Hydrometeor 4: 334-351.

Mitchell TD, Carter TR, Jones PD, Hulme M, New M (2004) A comprensive set of high resolution grids of monthly climate for Europe and Globe: and 16 scenarios (2001-2100), Tyndall, Centre for Climate Change Research, pp 55

Mölders N, Raabe A, Tetzlaff G (1996) A comparison of two strategies on land surface heterogeneity used in a mesoscale beta meteorological model. Tellus 48A: 733-749

Mölders N, Haferkorn U, Döring J, Kramm G (2003) Long-term numerical investigations on the water budget quantities predicted by the hydro-thermodynamic soil vegetation scheme (HTSVS) – Part I: Description of the model and impact of long-wave radiation, roots, snow and soil frost. Meteorol Atmos Phys 84: 115-135

Mölders N, Walsh JE (2004) Atmospheric response to soil-frost and snow in Alaska in March. Theor Appl Climatol 77: 77-105

Mölders N, Jankov M, Kramm G (2005) Application of Gaussian error propagation principles for theoretical assessment of model uncertainty in simulated soil processes caused by thermal and hydraulic parameters. J. Hydrometeorol 6: 1045-1062

Mölders N, Romanovsky VE (2006) Long-term evaluation of the Hydro-Thermodynamic Soil-Vegetation Scheme's frozen ground/permafrost component using observations at Barrow, Alaska. J Geophys Res 111: D04105, doi: 10.1029/2005JD005957

Montaldo N, and Albertson JD (2001) On the use of the force-restore SVAT model formulation for stratified soils. J. Hydrometeor. 2: 571-578.

Narapusetty B, Mölders N (2005) Evaluation of snow depth and soil temperatures predicted by the Hydro-Thermodynamic Soil-Vegetation Scheme (HTSVS) coupled with the PennState/NCAR Mesoscale Meteorological Model (MM5). J Appl Meteorol 44: 1827-1843

Narapusetty B, Mölders N (2006) Evaluation of the soil module of HTSVS by observations and a theoretically advanced numerical scheme. Mon Wea Rev 134: 2927-2942

Nijssen B, Schnur R, Lettenmaier DP (2001) Global retrospective estimation of soil moisture using the variable infiltration capacity Land Surface Model, 1980-93. J. Climate 14: 1790-1808

Niu G-Y, Yang Z-L, Dickinson RE, Gulden LE (2005) A simple TOPMODEL-based runoff parameterization (SIMTOP) for use in GCMs. J Geophys Res 110: D21106 doi:10.1029/2005JD006111

Nuss AW, Titley DW (1994) Use of multi-quadratic interpolation for meteorological objective analysis. Mon Wea Rev 122: 1611-1631

Oechel WC, Vourlitis GL, Hastings SJ, Zulueta RC, Hinzman L, Kane D (2000) Acclimation of ecosystem CO_2 exchange in the Alaskan Arctic in response to decadal climate warming. Nature 406: 978-981

Oleson KW, Dai Y, Bonan GB, Bosilovich M, Dickinson R, Dirmeyer P, Hoffman F, Houser P, Levis S, Niu GY, Thornton P, Vertenstein M, Yang ZL, Zeng X (2004) Technical description of Community Land Model (CLM). Technical Report NCAR/TN-461+STR, National Center for Atmospheric Research, Boulder, CO. 80307-3000, pp 174

Osterkamp TE (2003) A thermal history of permafrost in Alaska, Proceedings of the Eighth International Conference on Permafrost, 21–25 July 2003, Balkema Publishers, Zürich, Switzerland, pp 863–868

Pielke RA (2002) Mesoscale Meteorological Modeling, Academic Press, New York, 676 pp

Qu WQ, Henderson-Sellers A (1998) Comparing the scatter in PILPS off-line experiments with those in AMIP I coupled experiments. Global Planetary Change 19: 209-224

Rasch PJ, Kristjánsson JE (1998) A Comparison of the CCM3 Model Climate Using Diagnosed and Predicted Condensate Parameterizations. J Clim 11: 1587-1614

Romanovsky, VE, Sazonova TS, Balobaev VT, Shender NI, Sergueev DO (2007) Past and recent changes in air and permafrost temperatures in eastern Siberia. Global and Planetary Change (in press)

Rossow WB, Schiffer RA (1999) Advances in understanding clouds from ISCCP. Bull Amer Meteor Soc 80: 2261-2288

Sato N, Sellers PJ, Randall DA, Schneider EK, Shukla J, Kinter II JL, Hou Y-T, Albertazzi E (1989) Effects of implementing the Simple Biosphere Model in a general circulation model. J Atmos Sci 46: 2757-2782

Schlosser CA, Slater AG, Robock A, Pitman AJ, Vinnikov KY, Henderson-Sellers A , Speranaskaya NA, Mitchell K and the PILPS 2(D) contributors (2000): Simulation of a Boreal grass land hydrology at Valdai, Russia: PILPS phase 2(d). MonWea Rev 128: 301-321

Schmidt GA, Ruedy R, Hansen JE, Aleinov I, Bell N, Bauer M, Bauer S, Cairns B, Canuto V, Cheng Y, Del Genio A, Faluvegi G, Friend AD, Hall TM, Hu Y, Kelley M, Kiang NY, Koch D, Lacis AA, Lerner J, Lo KK, Miller RL., Nazarenko L, Oinas V, Perlwitz J, Rind D, Romanou A, Russell GL, Sato M, Shindell DT., Stone PH, Sun S, Tausnev N, Thresher D, Yao M-S (2006) Present-Day Atmospheric Simulations Using GISS ModelE: Comparison to in situ, satellite, and reanalysis data. J Climate 19: 153-192

Serreze MC, Walsh JE, Chapin FS, Osterkamp TE, Dyurgerov M, Romanovsky V, Oechel WC, Morison J, Zhang T, Barry RG (2000) Observational Evidence of Recent Change in the Northern High-Latitude Environment. Climatic Change 46: 159-207

Simmons AJ, Jones PD, da Costa Bechtold V, Beljaars ACM, Kållberg PW, Saarinen S, Uppala SM, Viterbo P, Wedi N (2004) Comparison of the trends and low frequency variability in CRU, ERA-40 and NCEP/NCAR analysis of surface air temperature. J. Geophys. Res. ,109, D24115, dio:10.1029/2004JD005306

Slater AG, Pitman AJ, Desborough CE (1998) The validation of snow parameterization designed for use in general circulation models. Int J Climatol 18: 595-617

Slater AG, Schlosser CA, Desborough CE, Pitman AJ, Henderson-Sellers A, Robock A, Vinnikov KY, Mitchell K, Boone A, Braden H, Chen F, Cox P, deRosany P, Dickinson E, Dai Y, Diuan Q, Entin J, Etchevers P, Gedney N, Gusev YM, Habets F, Kim J, Koren V, Kowalczyk E, Nasonova ON, Noilhan J, Schaake J, Shmakin AB, Smirnova TG, Verseghy D, Wetzel P, Xue Y, Yang Z-L, Zeng QC (2001) The Representation of Snow in Land Surface Scheme: Results from PILPS 2(d). Bull Amer Meteorol Soc 2: 7-25

Smith RD, Dukowicz JK, Malone RC (1992) Parallel ocean general circulation modeling. Physica D 60: 38-61

Stendel M, Christensen JH (2002), Impact of global warming on permafrost conditions in a coupled GCM. Geophys Res Lett 29: 1632, doi:10.1029/2001GL014345

Sugiura K, Yang D, Ohata T (2003) Systematic error aspects of gauge-measured solid precipitation in the Arctic, Barrow, Alaska. Geophys Res Lett 30: 1192, doi: 10.1029/2002GL015547

Sundquist H (1988) Parameterization of condensation and associated clouds in models for weather prediction and general circulation simulation. In Schlesinger ME (ed), Physically-based modeling and simulation of climate and climate change. Vol. 1, Kluwer Academic Publishers, Dordrecht, 433-461

Thiessen AH (1911) Precipitation for large areas. Mon Wea Rev 39: 1082-1084

Thompson SL, Pollard D (1995) A global climate model (GENSIS) with a land-surface transfer scheme (LSX). Part I: present climate simulation. J. Climate 8: 732-761

Uppala SM, Kållberg PW, Simmons AJ, Andrae U, Bechtold VD, Fiorino M, Gibson JK, Haseler J, Hernandez A, Kelly GA, Li X, Onogi K, Saarinen S, Sokka N, Allan RP, Andersson E, Arpe K, Balmaseda MA, Beljaars ACM, Berg L, Bidlot J, Bormann N, Caires S, Chevallier F, Dethof A, Dragosavac M, Fisher M, Fuentes M, Hagemann S, Hólm E, Hosking BJ, Isaksen L, Janssen PAEM, Jenne R, McNally AP, Mahfouf JF, Morcrette JJ, Rayner NA, Saunders RW, Simon P, Sterl A, Trenberth KE, Untch A, Vasiljevic D, Viterbo P, Woollen J (2005) The ERA-40 reanalysis. Q J R Meteorol Soc 131: 2961-3012

Viterbo P, Beljaars A, Mahouf J-F, Teixeira J (1999) The representation of soil moisture freezing and its impact on the stable boundary layer. Q J R Meteorol Soc 125: 2401-2426

Walsh JE, Chapman WL (1998) Arctic cloud-radiation-temperature associations in observational data and atmospheric reanalyses. J Climate 11: 3030-3045

Warrach K, Mengelkamp HT, and Raschke E (2001) Treatment of frozen soil and snow cover in the land surface model SEWAB. Theor Appl Climatol 69: 23-37

Yang D, Woo MK (1999) Representativeness of local snow data for large-scale hydrological investigations. Hydological Processes 13: 1977-1988

Yang D, Kane DL, Hinzman LD, Goodison BE, Metcalfe JR, Louie PYT, Leavesley GH, Emerson DG, Hanson CL (2000) An evaluation of the Wyoming gauge system for snowfall measurement. Water Resour Res 36: 2665-2678

Yang ZL, Pitman AJ, McAvaney B, Henderson-Sellers A (1995) The impact of implementing the bare essentials of surface transfer land surface scheme into the BMRC GCM. Climate Dynamics 11: 279-297

Zeng X, Shaikh M, Dai Y, Dickinson RE, Myneni R (2002) Coupling of the Common Land Model to the NCAR Community Climate Model. J Climate 15: 1832-1854

Zhang GJ, McFarlane NA (1995) Role of convective scale momentum transport in climate simulation. J Geophys Res 100:1417-1426

Zhang MH, Lin W, Bretherton CS, Hack JJ, Rasch PJ (2003) A modified formulation of fractional stratiform condensation rate in the NCAR community atmospheric model CAM2. J Geophys Res 108(D1):4035. DOI 10.1029/2002JD002523

Zhang T, Osterkamp TE, Stamnes K (1996) Influence of the depth hoar layer of the seasonal snow cover on the ground thermal regime. Water Resour Res 32: 2075-2086

Zhang T, Barry RG, Knowles K, Heginbottom JA, Brown J (1999) Statistics and characteristics of permafrost and ground ice distribution in the Northern Hemisphere. Polar Geography 23: 132-154

Zhang T, Barry RG, Gilichinsky D, Bykhovets SS, Sorokovikov VA, Ye J (2001) An amplified signal of climatic change in soil temperatures during the last century at Irkutsk, Russia. Climatic Change 49: 41-76

Zhang T, Frauenfeld OW, Serreze MC, Etringer A, Oelke C, McCreight J, Barry RG, Gilichinsky D, Yang D, Ye H, Ling F, Chudinova S (2005) Spatial and temporal variability in active layer thickness over the Russian Arctic drainage basin. J Geophys Res 110D 16101: doi:10.1029/2004JD005642

Zhuang Q, Romanovsky VE, McGuire AD (2001) Incorporation of a permafrost model into a large-scale ecosystem model: Evaluation of temporal and spatial scaling issues in simulating soil thermal dynamics. J Geophys Res 106D: 33,649-33,670

Table 4.1 Annual average of simulated and observed quantities, BIAS, RMSE, SDE, and correlation of various quantities for 1971-2000. Values in brackets refer to 1961-1990 and 1951-1980, respectively. Note that values for snow depth refer to 1978 to 2003. In the case of near-surface air temperatures, values in brackets are for the second climatology. For the cloud fraction values are valid for 1984-2004.

Quantity	Unit	Annual simulated average	Annual observed average	BIAS	RMSE	SDE	Correlation coefficients
Soil temperature	K	275.9 (275.7, 275.7)	273.4 (273.3, 273.1)	2.5 (2.4, 2.6)	4.4 (4.4, 4.5)	2.7 (2.7, 2.6)	0.865 (0.867, 0.871)
Near-surface temperature	K	271.9 (271.5)	272.9 (272.7)	-1.0 (-1.1)	3.0 (3.0)	2.5 (2.6)	0.954 (0.954)
Cloud fraction	%	66.66	66.33	0.32	0.52	0.76	0.539
Precipitation	mm/ mon	44.0 (43.7, 43.7)	36.4 (36.6, 36.8)	7.6 (7.1, 6.9)	19.9 (20.1, 20.4)	15.9 (16.3, 16.6)	0.752 (0.746, 0.743)
Snow depth	m	0.194	0.154	0.040	0.096	0.087	0.541

Table 4.2 Correlation coefficients of difference between simulated and observed soil temperature with the difference between simulated and reanalyzed near-surface air temperature for the second and third climatology for winter and summer. Note that all the correlations are significant at the 95% confidence level.

Climatology	Season	0.2 m	0.4 m	0.8 m	1.6 m	3.2m
Second	Winter	0.425	0.429	0.404	0.53	0.329
	Summer	0.720	0.685	0.582	0.317	0.366
Third	Winter	0.373	0.389	0.382	0.507	0.352
	Summer	0.715	0.688	0.581	0.299	0.337

Table 4.3 Correlation coefficients of difference between simulated and observed soil temperature with the difference between simulated and observed precipitation for all three climatologies for winter and summer. Note that significant correlations at the 95% confidence level are given in bold.

Climatology	Season	0.2 m	0.4 m	0.8 m	1.6 m	3.2 m
First climatolology	Winter	-0.048	-0.026	-0.071	0.074	**-0.155**
	Summer	**-0.445**	**-0.369**	**-0.222**	**-0.048**	0.014
Second climatolology	Winter	-0.01	0.016	-0.034	**0.103**	**-0.106**
	Summer	**-0.451**	**-0.374**	**-0.186**	**-0.0242**	0.045
Third climatolology	Winter	-0.029	-0.002	-0.050	**0.103**	**-0.111**
	Summer	**-0.504**	**-0.377**	**-0.242**	-0.063	0.029

Table 4.4 Correlation coefficients of difference between simulated and observed soil temperature with the difference between simulated and observed snow depth for 1979-2000. Note that all the correlations are significant at the 95% confidence level.

Months	0.2 m	0.4 m	0.8 m	1.6 m	3.2 m
November	0.615	0.560	0.438	-0.038	0.025
December	0.579	0.561	0.487	0.225	0.232
January	0.559	0.538	0.485	0.378	0.337
February	0.526	0.504	0.456	0.365	0.280
March	0.457	0.447	0.413	0.365	0.225

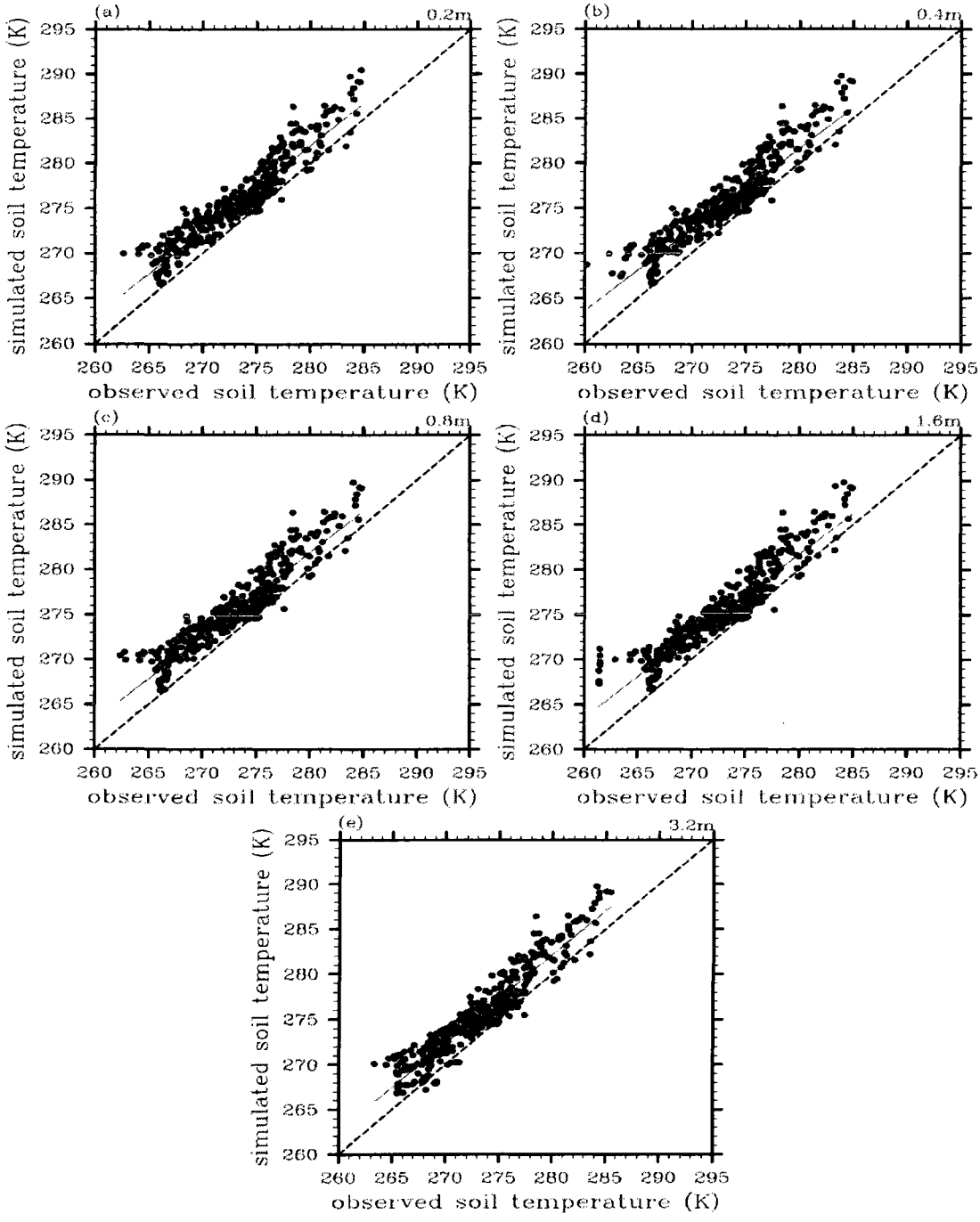


Fig. 4.1 Comparison of annually-averaged simulated and observed soil temperatures at (a) 0.2 m, (b) 0.4 m, (c) 0.8 m, (d) 1.6 m, and (e) 3.2 m depth for the first climatology. The 1:1 line (dashed) and the regression line (thin solid) are superimposed. Note that scatter plots for the second and third climatology look similar.

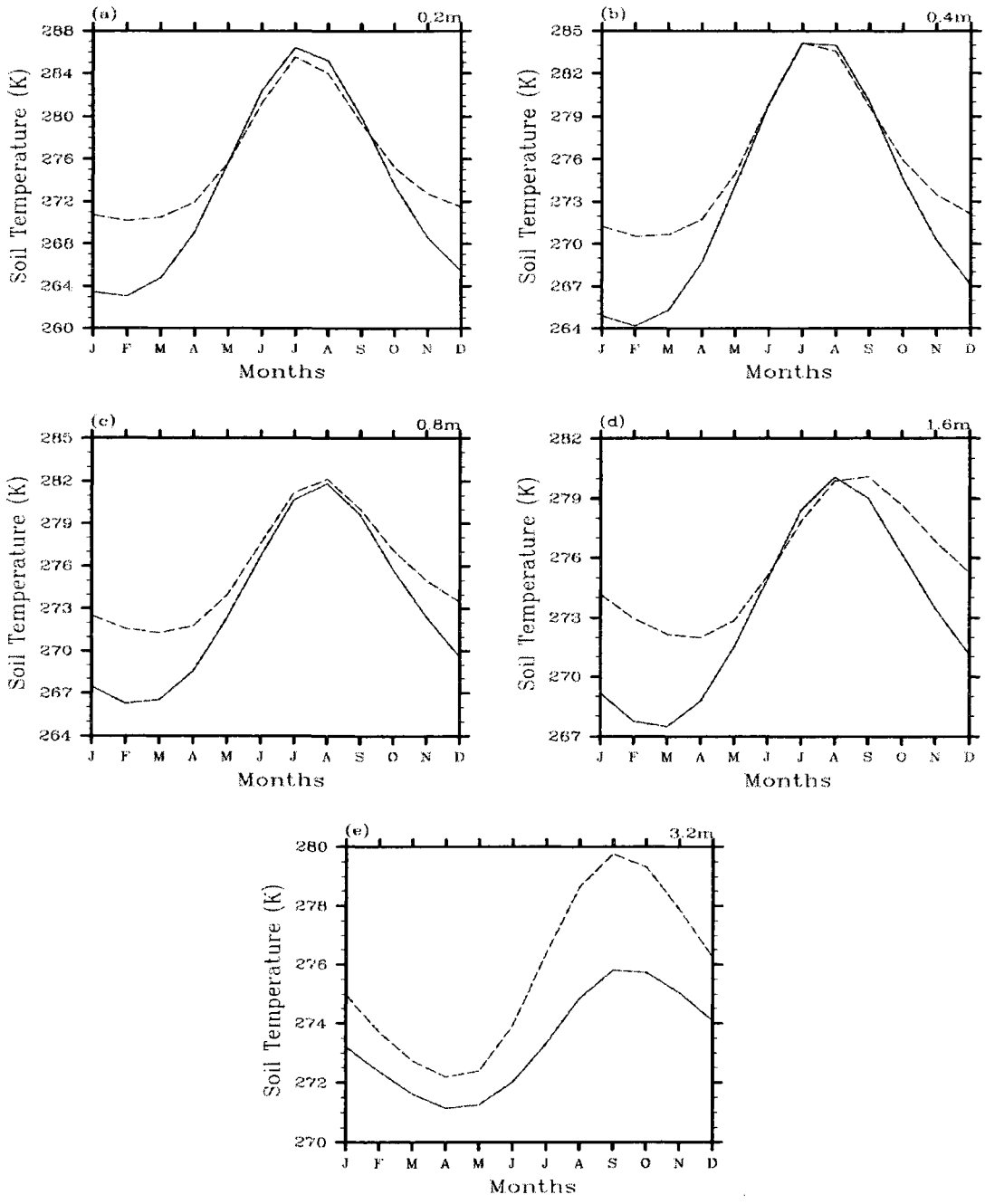


Fig. 4.2 Temporal behavior of monthly domain-averaged simulated (dashed) and observed (solid) soil temperatures at (a) 0.2 m, (b) 0.4 m, (c) 0.8 m, (d) 1.6 m, and (e) 3.2 m depth for the first climatology. Curves of the second and third climatology look similar (therefore not shown). Note that the scaling of x- and y-axes differs among parts of the panel.

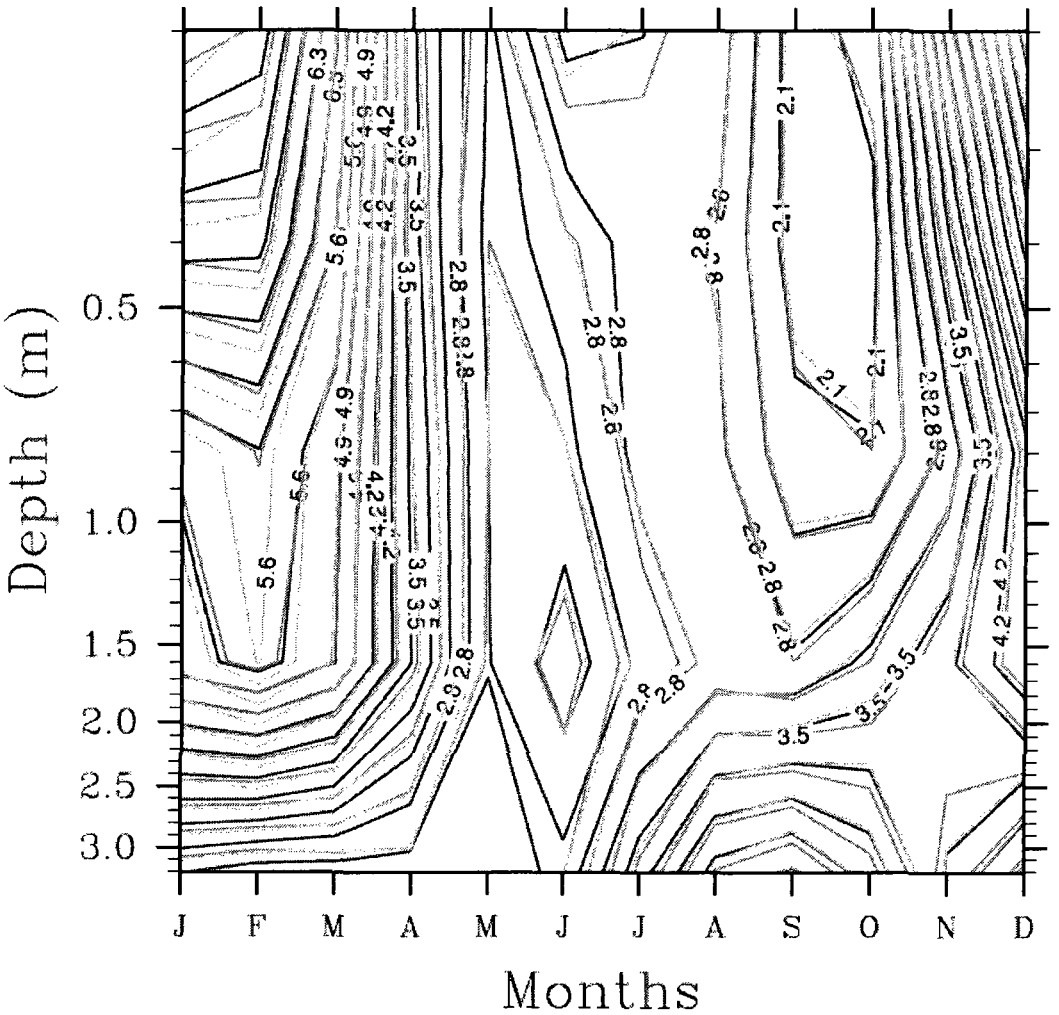


Fig. 4.3 Contour plot of the temporal behavior of RMSEs with depth for the first climatology (black), second climatology (gray), and third climatology (light gray).

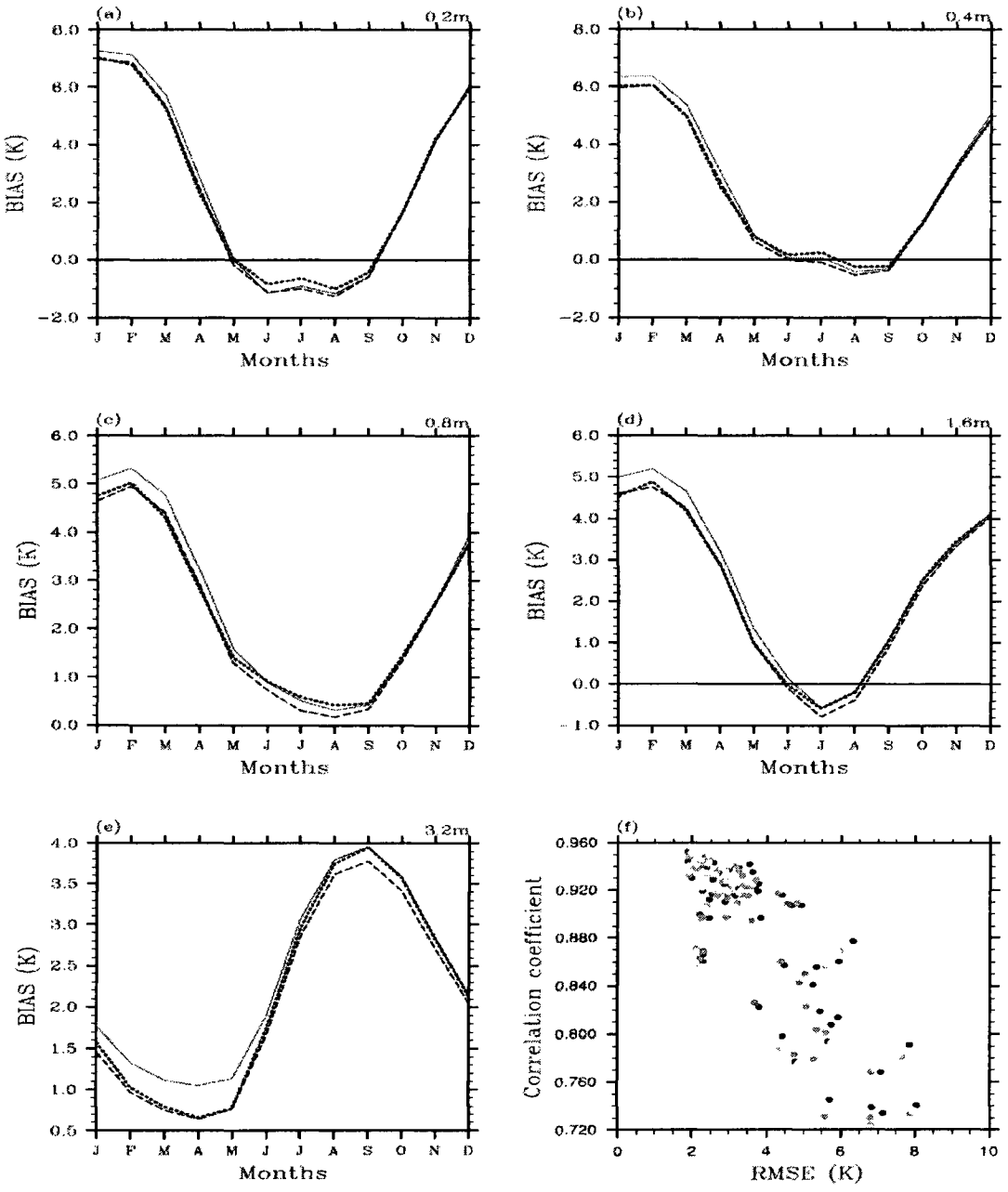


Fig. 4.4 Temporal behavior of monthly domain-average BIASes of soil temperatures at (a) 0.2 m, (b) 0.4 m, (c) 0.8 m, (d) 1.6 m, and (e) 3.2 m depth. The solid, dashed, and dotted lines represent the first, second, and third climatology, respectively. (f) RMSE vs. correlation coefficient for first (black), second (dark gray), and third (light gray) climatology for all 12 months at 0.2 m , 0.4 m, 0.8 m, 1.6 m, and 3.2 m depth. Note that the scaling of x- and y-axes differs among parts of the panel.

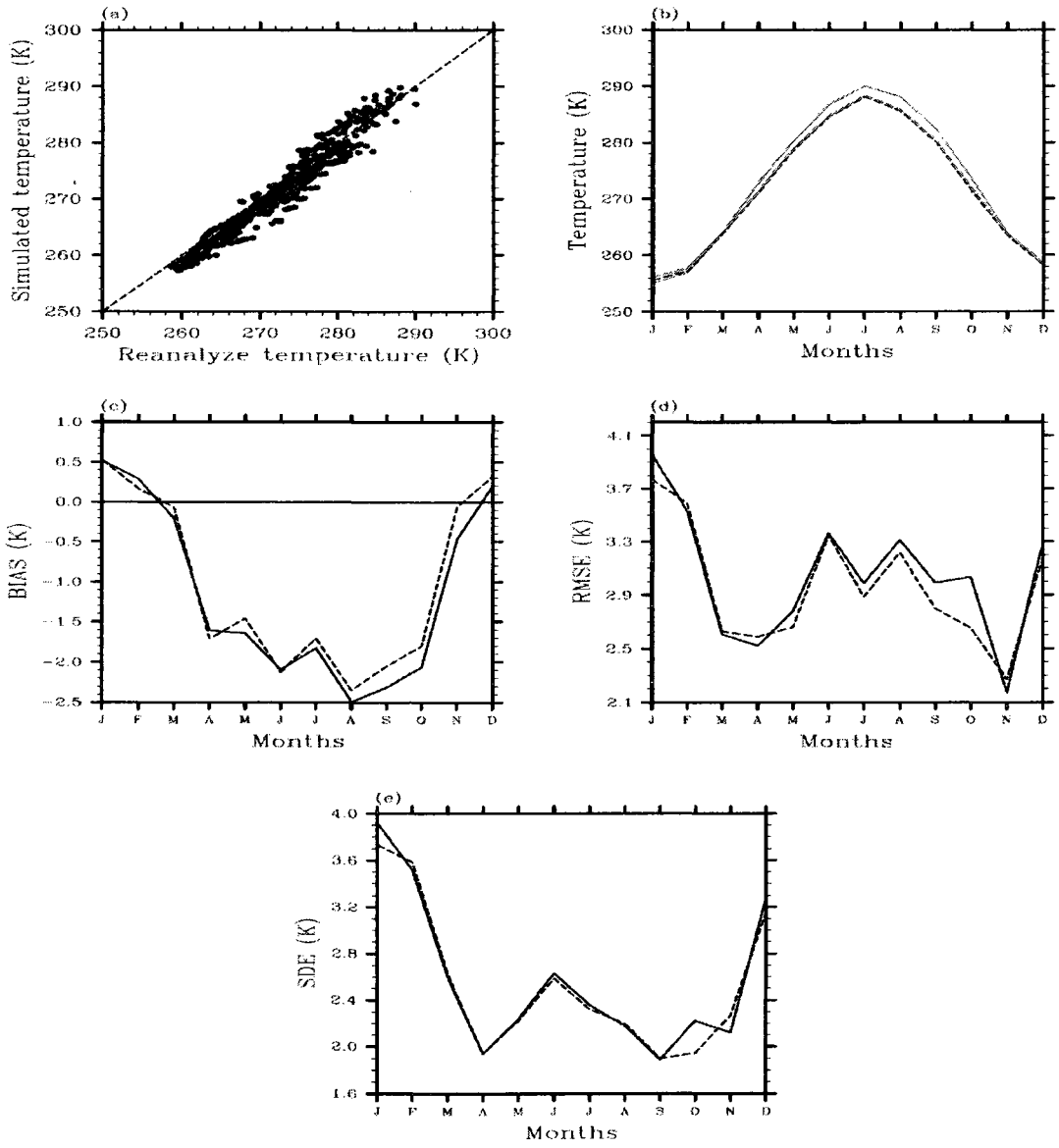


Fig 4.5 (a) Scatter plot of annually-average reanalyzed near-surface air temperature and simulated near-surface air temperature for second climatology with 1:1 line superimposed and the temporal behavior of (b) domain-averaged near-surface temperature climatology derived from the CCSM simulation (dashed) and ERA40 data (solid) with the second and third climatology shown as dark and light lines, respectively, (c) BIAS, (d) RMSE, and (e) SDE between reanalyzed near-surface air temperature and simulated near-surface air temperature for second (solid line) and third (dashed) climatology. Note that the scatter-plots for the third climatology look similar and ERA40 data do not cover the first climatology. The scaling of x- and y-axes differs among parts of the panel.

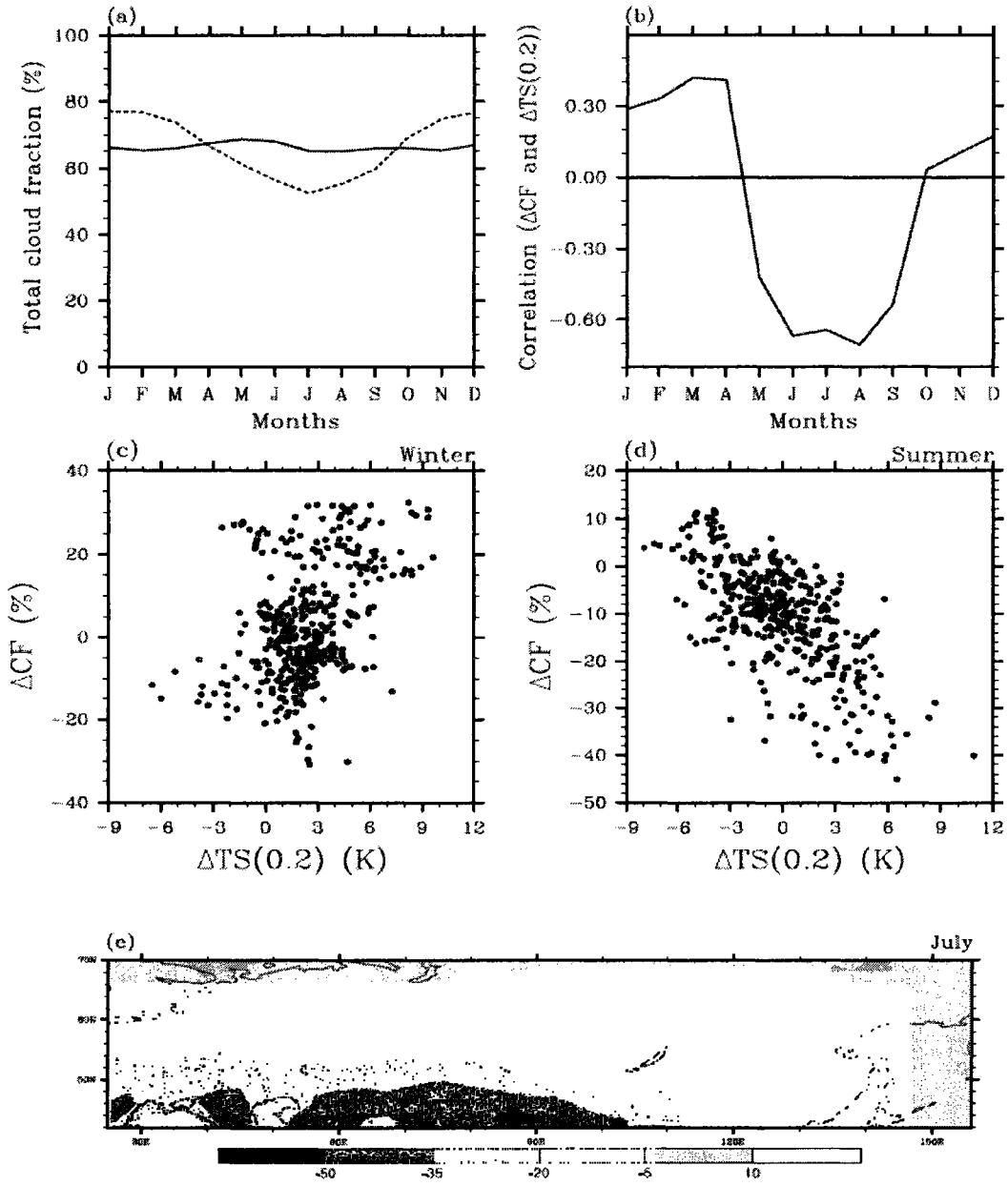


Fig. 4.6 Evaluation of cloud fraction (a) annual cycle of domain-averaged CCSM-simulated cloud fraction (dotted line) and ISCCP cloud fraction (solid line), (b) temporal evolution of correlation of differences between simulated and observed cloud fraction, ΔCF with differences between simulated and observed soil temperature at 0.2 m depth, ΔTS(0.2); scatter plot of differences between simulated and observed soil temperatures, ΔTS(0.2) and differences between simulated and observed cloud fraction, ΔCF for (c) winter and (d) summer. (e) Horizontal distribution of differences between simulated and observed cloud fraction over Russia in July.

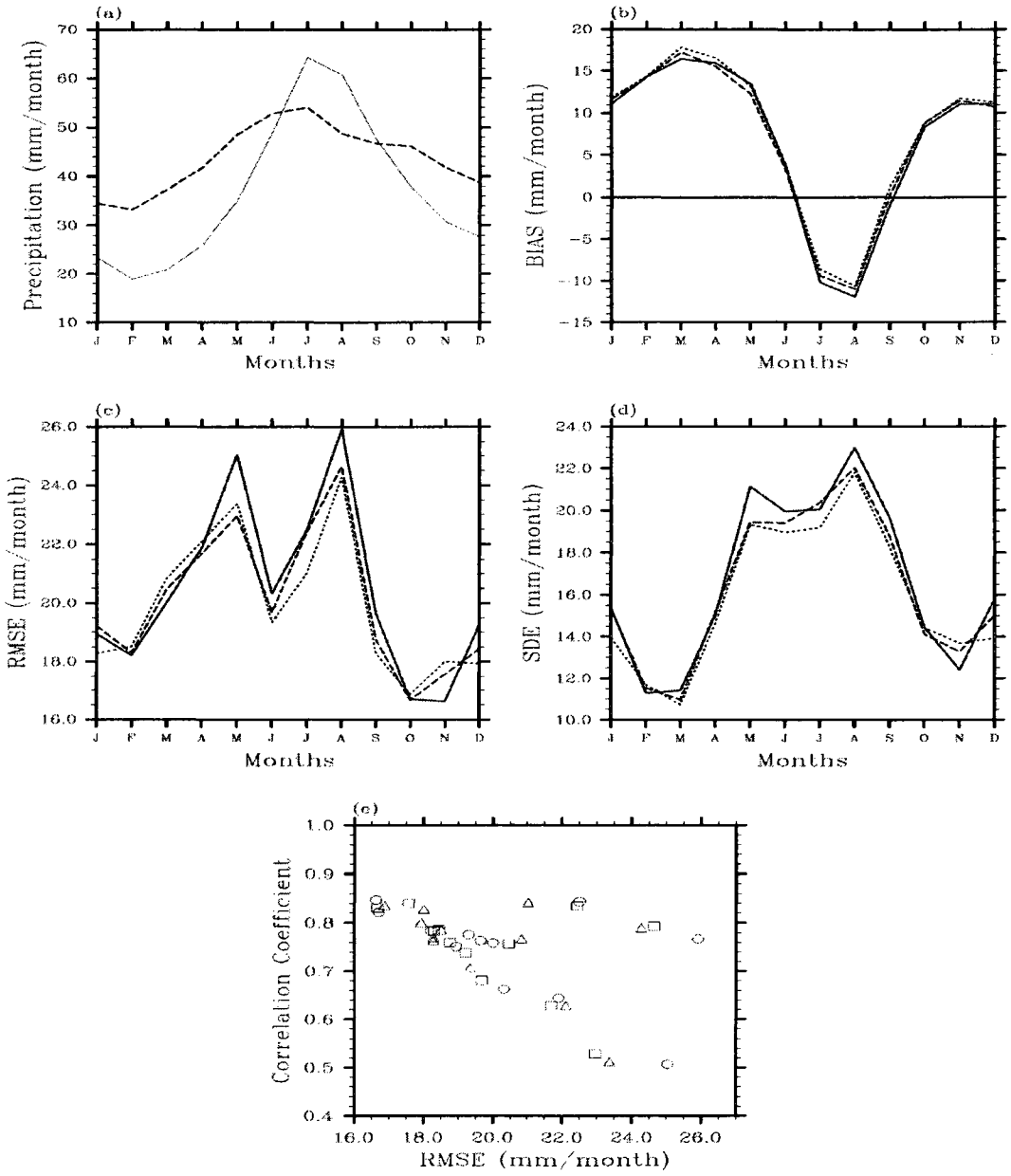


Fig. 4.7 Temporal behavior of (a) domain-averaged simulated (dashed line) and observed (solid line) precipitation for first climatology. Curves for the second and third climatology look similar (therefore not shown). (b) BIAS, (c) RMSE, and (d) SDE between observed and simulated precipitation for the first (dotted), second (solid), and third (dashed) climatology. (e) Comparison of RMSE and correlation coefficient between simulated and observed precipitation for first (circle), second (square), and third (triangle) climatology.

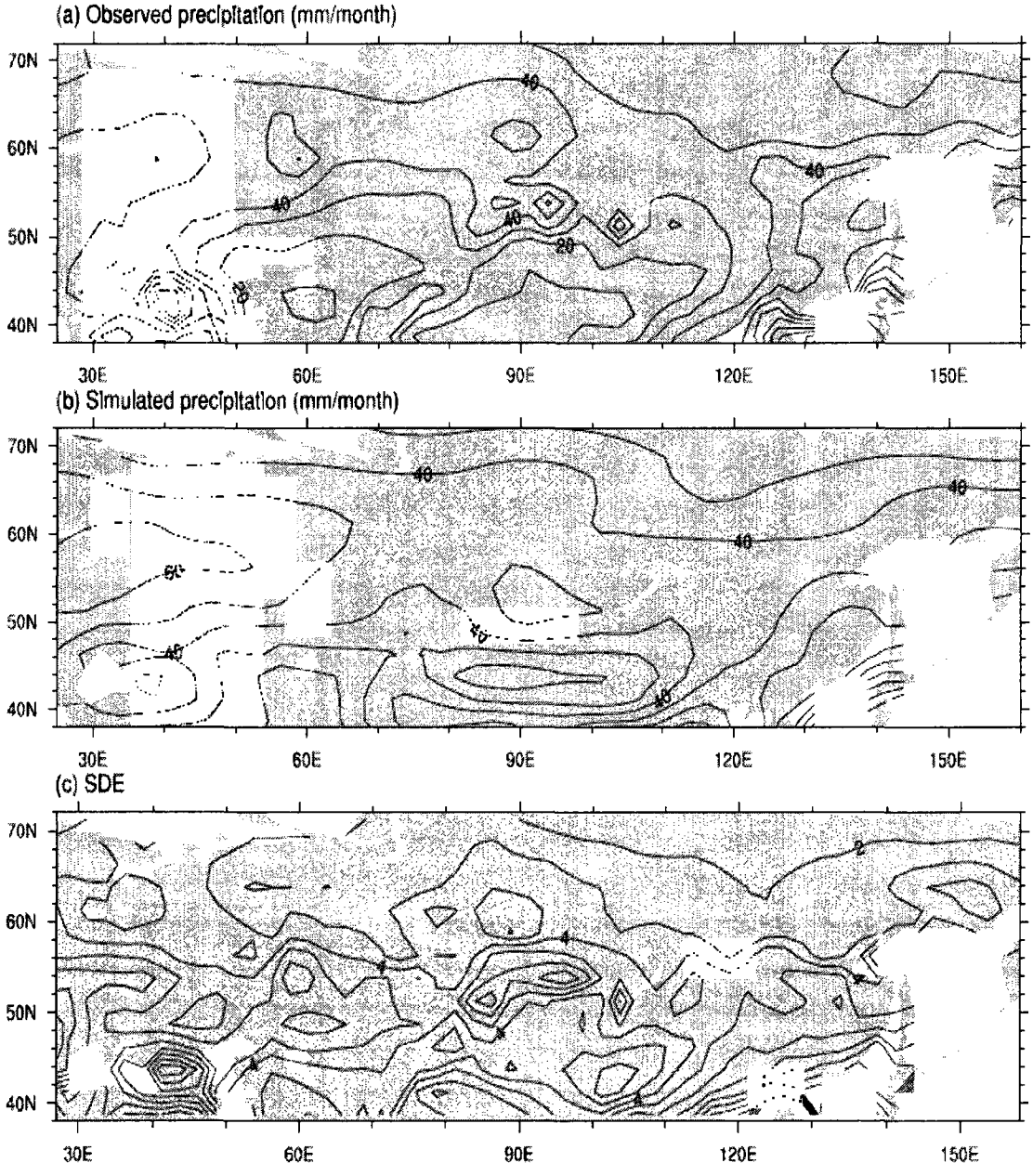


Fig. 4.8 Horizontal distribution of (a) observed and (b) simulated 30-year annual average precipitation and (c) SDE between simulated and observed precipitation for the first climatology over Russia. Note that values are shown only for grid-cells with land because no observations are available for ocean. Plots for other climatologies look similar (therefore not shown).

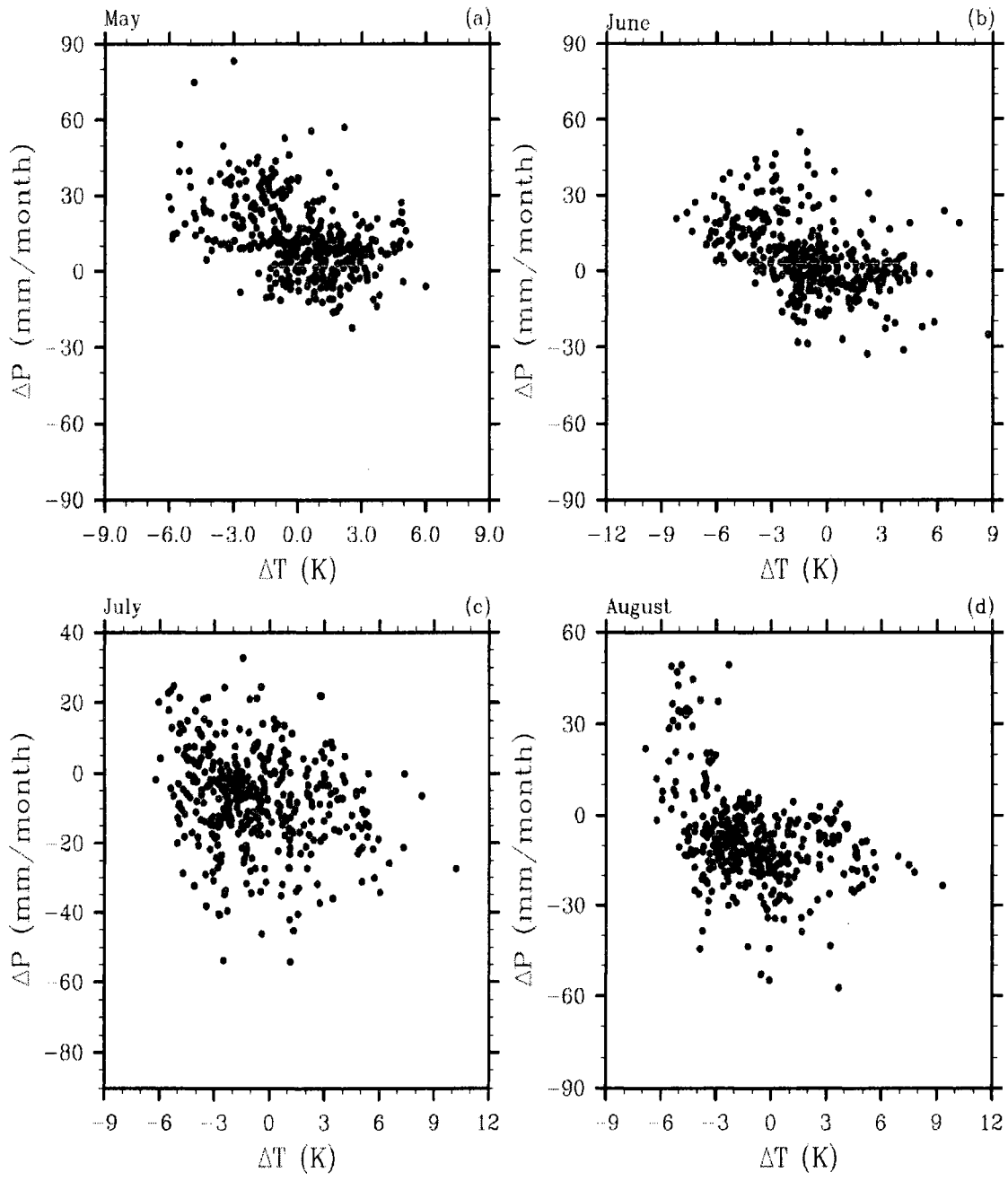


Fig. 4.9 Difference between simulated and observed soil temperatures ΔT at 0.2 m depth vs. difference in simulated and observed precipitation ΔP for the first climatology for (a) May, (b) June, (c) July, and (d) August. Scatter-plots for second and third climatology look similar. Note that the scaling of x- and y-axes differs among parts of the panel.

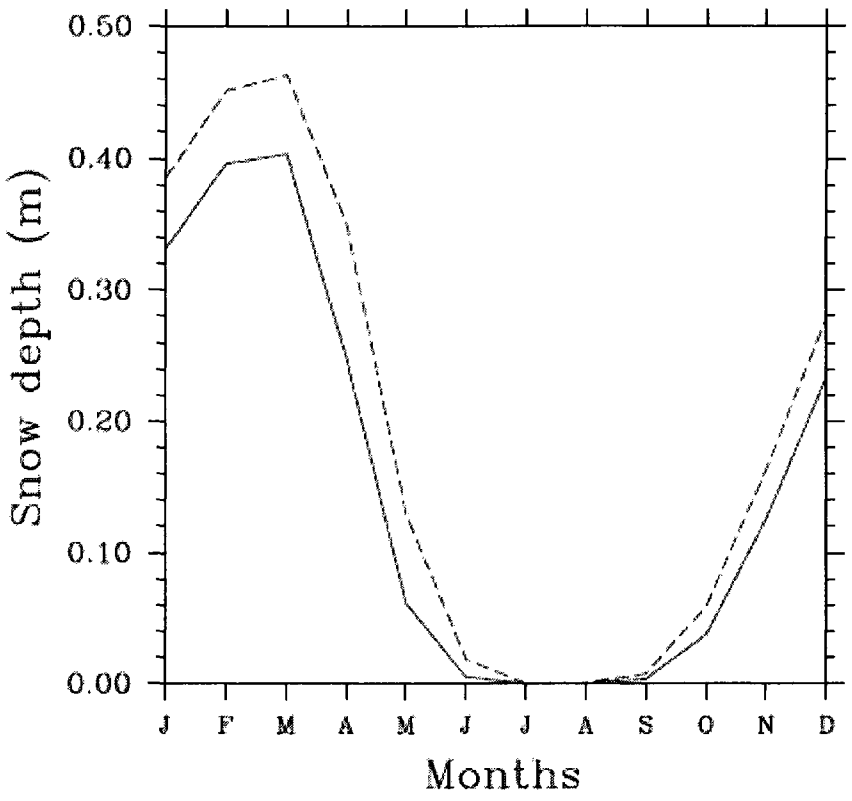


Fig. 4.10 Temporal behavior of domain-averaged simulated snow depths (dashed) and the 1979-2003 observed snow depths (solid).

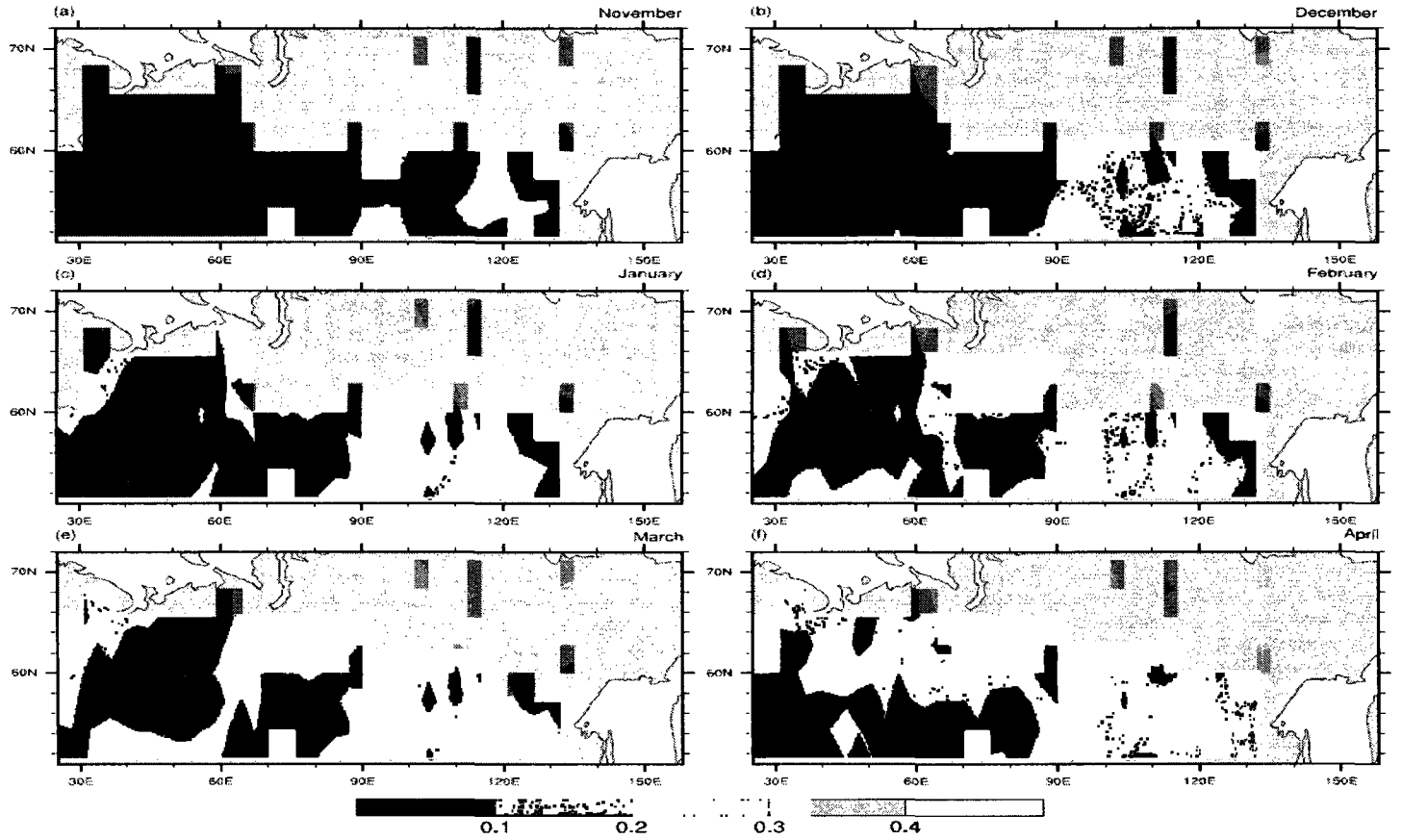


Fig. 4.11 Horizontal distribution of RMSE between simulated and observed snow depth for (a) November, (b) December, (c) January, (d) February, (e) March, and (f) April.

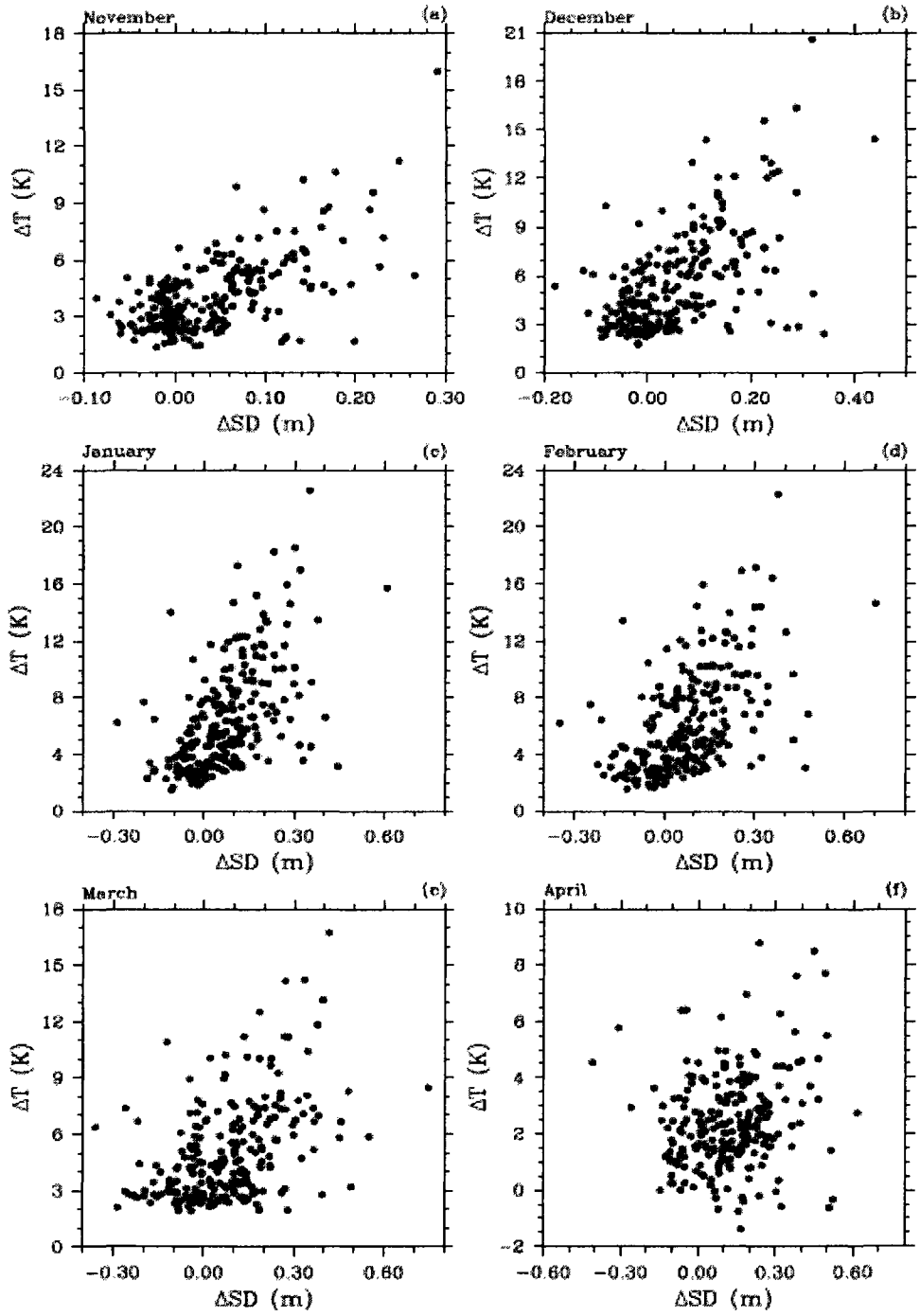


Fig. 4.12 Difference between simulated and observed soil temperatures ΔT at 0.2 m depth vs. difference in simulated and observed snow depth ΔSD for 1979-2003 for (a) November, (b) December, (c) January, (d) February, (e) March, and (f) April. Note that x- and y-axis scaling varies among the parts of the panel.

Chapter 5 Theoretical assessment of uncertainty in regional averages due to network density and design[‡]

Abstract

Weather Research and Forecasting (WRF) model simulations are performed over Russia for July and December 2005, 2006 and 2007 to create a "dataset" to assess the impact of network density and design on regional averages. Based on the values at all WRF grid-points regional averages for various quantities are calculated for $2.8^{\circ} \times 2.8^{\circ}$ areas as the "reference". Regional averages determined based on 40 artificial networks and 411 "sites" that correspond to the locations of a real network, are compared with the reference regional averages. The 40 networks encompass ten networks of 500, 400, 200, or 100 different randomly taken WRF-grid-points as "sites".

The real network's site distribution misrepresents the landscape. This misrepresentation leads to errors in regional averages that show geographical and temporal trends for most quantities: errors are lower over shores of large lakes than coasts and lowest over flatland followed by low and high mountain ranges; offsets in timing occur during frontal passages when several sites are passed at nearly the same time. Generally, the real network underestimates regional averages of sea-level pressure, wind-speed, and precipitation over Russia up to 4.8 hPa (4.8 hPa), 0.7 m/s (0.5 m/s), and 0.2 mm/d (0.5 mm/d), and overestimates regional averages of 2-m temperature, downward shortwave radiation and soil-temperature over Russia up to 1.9 K (1.4 K), 19 Wm^{-2}, (14 Wm^{-2}), and 1.5 K (1.8 K) in July (December). The low density of the ten 100-sites-networks causes difficulties for sea-level pressure. Regional averages obtained from the 30 networks with 200 or more randomly distributed sites represent the reference regional averages, trends and variability for all quantities well.

[‡]PaiMazumder D, Mölders N (2009) Theoretical assessment of uncertainty in regional averages due to network density and design. J Appl Meteor Climatol (in press)

5.1 Introduction

Appropriate meteorological networks are among the important pre-requisites to evaluate numerical weather prediction and climate models of various scales, to determine representative regional averages for climatology and to identify climate changes. Mesoscale-γ/β models are typically evaluated by assuming that measurements at a site are representative for the grid-cell the site falls into (Chase et al. 1996, Zhong et al. 2005, Mölders and Kramm 2007). This assumption cannot be made for General Circulation Models (GCMs) because here fluxes and state variables represent volume and area averages of several 100 square-kilometers in horizontal and several decameters in vertical direction. Furthermore, often several sites may exist within the area represented by a GCM grid-cell making a comparison like in mesoscale modeling ambiguous. Therefore, in climate modeling, it has become common practice for evaluation purposes to use interpolation methods and/or calculate regional averages to produce gridded data for areas of the size of GCM grid-cells based on the available measurements (Palutikof et al. 1997, Bauer et al. 2002, Li et al. 2008, PaiMazumder et al. 2008). However, doing so bears uncertainty from the interpolation methods and observations.

Uncertainty in gridded regional averages has been examined with respect to the interpolation methods for precipitation, radiation, air pollutants and meteorological state variables (Shaw and Lynn 1972, Creutin and Obled 1982, Court and Bare 1984, Lebel et al. 1987, Lindley and Walsh 2004, Luo et al. 2008). Major findings were that (1) any interpolation technique causes uncertainty in regional averages; (2) the choice of interpolation methods should depend on the nature of the region and available data (type, amount); and (3) some interpolation methods are not well-suited for regions with strong systematic variations unless the site density is high and the sites are well-distributed over the region. Common conclusions were that (1) optimal interpolation methods provide the best results for regional precipitation because they include the spatial correlation structure of precipitation; (2) kriging (a statistical technique based on auto-correlation to interpolate the variables of a random field on a grid from data at observational sites) provides the best results for spatial interpolation of pollutant concentrations, precipitation

and temperature (Tabios and Salas 1985, Lefohn et al. 1987, Holdaway 1996, Phillips et al. 1997, Ninyerola et al. 2000, Jeffrey et al. 2001, Lindley and Walsh 2004); and (3) even with an optimal interpolation method regional precipitation and temperature averages can be biased by observers, poorly-sited stations, network design and/or using data that were originally collected for other purposes. Under-reporting and observers' preference to report precipitation values divisible by 5 and/or 10, for instance, can cause bias in regional precipitation averages (Daily et al. 2007). Near-surface temperatures obtained from poorly- and inhomogeneouly-sited stations vary stronger in comparison to North American Regional Reanalysis (Mesinger et al. 2006) than well-sited stations (Pielke et al. 2007). Changes in site location or network density alter the topography, latitude and elevation represented by the network, difference in sensors and their exposures and site exposure to cold air; these changes affect air-temperature, minimum and maximum temperature measured (Robeson and Doty 2005, Peterson 2006) with consequences for regional averages calculated by means of these sites. Not-representing the topography by the network may cause systematic bias in regional average precipitation because the measurements are made at different terrain heights than those that would reflect the region (Groisman et al. 1991, Groisman and Legates 1994). Network density may also affect regional precipitation averages (Frei and Schär 1998, Tsintikidis et al. 2002); high-density networks are more likely to capture locally high precipitation rates than coarse networks (St.-Hilarie et al. 2003). Especially, if a huge fraction of precipitation stems from convection, regional averages determined from high-density networks may be more accurate than those of coarse networks. Uncertainty due to networks becomes particularly problematic in remote areas, where networks are often designed with accessibility and ease of maintenance in mind. Consequently, these networks follow major haul-ways and are not randomly distributed. Accuracy and reliability of long time-series of gridded data compiled from all available stations may be influenced by urbanization, land-cover changes, moving, shutting down or adding sites, errors in digitizing old paper records, the procedure of filling missing data and, to a certain degree, on the applied interpolation algorithms (Mitchell et al. 2004).

Today, gridded data from networks with long time series are often used for GCM evaluation, i.e. some of these networks already existed before GCMs became available. The gridded soil-temperature data (Zhang et al. 2001) used by PaiMazumder et al. (2008) for GCM evaluation, for instance, stem from long-term agricultural monitoring stations; the lysimeter-data used by Mölders et al. (2003) for evaluation of the water budget of the Hydro-Thermodynamic Soil Vegetation Scheme (Kramm et al. 1996) were originally collected to assess ground-water recharge. Since it takes decades to sample long time series one has to put aside what purpose a long-term monitoring network was designed for (Goody et al. 2002), but assess what one can scientifically meaningful do with its data, how limited they are and what uncertainty they bear. Examining this for a real network is one of the goals of this study.

The fluxes that change the state variables in the system Earth-atmosphere depend on those states (Entekhabi and Brubaker 1995). Because of the non-linear dynamical modes of variability and statistical signatures related to these interactions results found for the impact of network design and/or density on regional averages of precipitation, concentration and temperature cannot be generalized for other state variables and fluxes. Therefore, to separate GCM weaknesses in representing complex processes from uncertainty due to observation-derived climatology, it is essential to understand the potential impact of network density and/or site distribution on gridded regional averages for a broad variety of state variables and fluxes.

In our case study, we assess this uncertainty in gridded regional averages for a large variety of quantities for July and December 2005, 2006 and 2007 over Russia (Fig. 5.1). In doing so, we use the Weather Research and Forecasting (WRF; Skamarock et al. 2005) model to create a "reference dataset". Regional averages for $2.8^{\circ} \times 2.8^{\circ}$, a common size of GCM grid-cells for the last decade's era of climate modeling, are determined from the WRF-predicted values. These "reference" regional averages are compared with regional averages derived from 40 different artificial networks, ten each of four different densities (500, 400, 200, and 100 sites) with randomly distributed sites and a non-randomly distributed network (Fig. 5.2). The non-randomly distributed network is based on the site

locations of a real network that has over 50 years of soil-temperature data for which it is frequently used in climate model evaluation and climatological studies (Zhang et al. 2001, Romanovsky et al. 2007, PaiMazumder et al. 2008). Since it has become a classical long-term dataset that is widely used because of its consistency and length, the sites of this dataset are chosen. Advantages of using these sites in the investigation are twofold: the investigations will be provided with a better sense of errors caused by this network and may help the scientific community to assess difference between their simulations and the gridded soil-temperature observations. July and December are chosen because these are the months with the greatest discrepancies between GCM simulated and observed soil-temperatures (PaiMazumder et al. 2008) and locally great changes with time (Romanovsky et al. 2007). The advantage of using WRF-generated values over randomly generated values is that the former not only permits us to assess the degree of potential uncertainty in gridded climatology related to network design and/or density, but also provides hints where/when regional averages of gridded data may be more or less reliable.

5.2 Experimental design

5.2.1 Reference dataset

WRF simulations are performed with the model-setup given in Table 5.1 for July and December 2005, 2006, and 2007. The model domain encompasses 70×150 grid-points over Russia (Fig. 5.1) with a 50km-grid-increment and 31 vertical layers from the surface to 50 hPa and six layers in the soil; in the presence of snow, five snow layers are considered. Simulations start daily at 1800 UT for 30 h of integration. We discard the first six hours of each simulation as spin-up time. National Centers for Environmental Prediction $1^{\circ} \times 1^{\circ}$ and 6h-resolution global final analyses serve as initial and boundary conditions.

The WRF-simulated quantities are assumed to be "measurements" from an optimal, dense, and spatially equally distributed "observational network" referred to as reference. Regional averages of sea-level pressure (SLP), 10-m wind-speed, 2-m temperature,

minimum and maximum temperature, precipitation, relative humidity, shortwave and long-wave radiation and soil-temperature are determined for areas of $2.8^\circ \times 2.8^\circ$ using all WRF-simulated values that fall into these areas and a Cressman-type method as described in PaiMazumder et al. (2008). There are 637 (13×49) grid-cells of $2.8^\circ \times 2.8^\circ$ areas in the model domain (39.5°N-73°N, 22°E-157°E). Regional averages calculated for these 637 $2.8^\circ \times 2.8^\circ$ areas are called "reference" hereafter.

5.2.2 Networks

Forty networks, ten with 500, 400, 200, and 100 sites each, called 500-, 400-, 200- and 100-sites-networks hereafter, are assumed with 500, 400, 200, and 100 randomly taken WRF-grid-points as "sites" (Fig. 5.2). These 40 networks are chosen by using a random number generator over land grid-cells only. Regional averages for $2.8^\circ \times 2.8^\circ$ areas are calculated based on these networks. If several "sites" fall within an $2.8^\circ \times 2.8^\circ$ area the same method as for determining the reference regional averages is used to calculate the regional average. In the following, the term "regional averages" refers to averages for $2.8^\circ \times 2.8^\circ$ areas.

Furthermore, regional averages are determined for 411 sites of an existing Russian network (Fig. 5.2) called real network hereafter. In doing so, the WRF-variables of the grid-cells the sites fall into are assumed to be the "observations". There are up to 2, 5, 5, 6 and 8 WRF-sites within the $2.8^\circ \times 2.8^\circ$ areas of the 100, 200, 400, 500-sites-networks and the real network. These values represent the maximum number of WRF-sites within the $2.8^\circ \times 2.8^\circ$ areas occurring in the ten artificial networks of different densities and the real network.

The regional averages determined based on the "sites" of the real and 40 artificial networks are compared with the "reference" to assess the contribution of network density and/or design to uncertainty in regional averages of gridded data. To reduce uncertainty from interpolations in areas that have no "site" at all, we only discuss (and show) results for $2.8^\circ \times 2.8^\circ$ areas with at least one "site". This procedure leaves us with up to 310, 277,

168, 89 and 231 regional averages of $2.8^{\circ} \times 2.8^{\circ}$ areas per time-period investigated for the 500, 400, 200, 100-sites-networks in all ten cases and the real network, respectively.

5.2.3 Analysis method

We calculate the difference between the reference regional averages and regional averages derived from the various networks to quantify the uncertainty in the gridded data/regional averages caused by network density and/or design. Any regional average calculated from a network will be considered as being in excellent agreement with the reference regional average if it falls within the reference value plus/minus the accuracy of routine measurements given in Table 5.2.

We determine performance measures (Table 5.3) to identify reasons for discrepancies in regional averages. In our experimental design, bias indicates systematic errors in regional averages caused by differences in physical and/or geometric factors between the landscape represented by a network and the reference landscape (terrain elevation, vegetation-type, vegetation-fraction, soil-type, etc.). The standard deviation of error (SDE) quantifies random error when the bias is removed. In any actual observational network, measurements have random errors. In our experimental design, random error may stem, among other things, from initialization and boundary conditions. At the bottom of the soil model, for instance, soil-temperature and moisture vary spatially, but not with time. Taking a "site" just one WRF-grid-point apart may mean a different boundary condition with impacts on soil-temperature and moisture. Root-mean-square-error (RMSE) assesses the overall success of a network in capturing the reference regional average and avoids positive and negative differences canceling each other out. To measure the strength of the various networks in capturing trends and/or phases of regional averages, correlation-skill scores (r) between the regional averages derived for the various networks and the reference regional averages are determined. To determine the overall relative degree to which the regional averages derived from various networks approach the reference regional averages, Willmott's index of agreement (Willmott 1984, Cannon and Whitfield 2002) between the regional averages derived from the reference

network and the various artificial networks is calculated. Willmott's index of agreement (WIA) ranges from 0 (complete disagreement) to 1 (perfect agreement).

For all networks for all $2.8^{\circ} \times 2.8^{\circ}$ areas with at least one "site", performance measures are calculated based on hourly values. To examine shifts in regional averages related to the networks' representation of timing of events (frontal passage, heating/cooling, convection) we examine the averages and measures for the domain at large (Table 5.4).

5.3 Results

5.3.1 Representation of landscape

WRF uses the strategy of dominant land-cover. This strategy assumes that the land-cover type dominating in a grid-cell is representative for the entire grid-cell and can be used to calculate the exchange of momentum, heat and moisture at the surface-atmosphere interface. Consequently, the landscape assumed in WRF is more homogeneous and much simpler than a natural landscape (Avissar and Pielke 1989). This WRF-assumed reference landscape is a mix of broadleaf and coniferous forest, wetlands and tundra (Fig. 5.3a) partly underlain by warm permafrost. The 100, 200, 400 and 500-sites-networks in all their ten setups represent the frequency of occurrence of the various land-cover types in the reference landscape within $\pm 5\,\%$. The real network significantly ($95\,\%$ or higher confidence level) overestimates the fraction of mixed forest (MF), savanna (SV) and needle-leaf evergreen forest (NEF), and underestimates the extension of water-bodies (WB) by $10\,\%$. These misrepresentations of land-cover may cause some uncertainty in regional averages of energy balance components, 2-m temperature, wind-speed, relative humidity and precipitation derived from the real network (section 5.3.5). All other land-cover types are within $\pm 5\,\%$ of the fraction found in the reference landscape (Fig. 5.3a). Note that the real network was originally designed to monitor conditions in agriculturally used land (cf. Zhang et al. 2001); the landscape considered by WRF, however, has a variety of land-cover types for the WRF-grid-cells the sites of the real network fall into. Thus, any discrepancies found for the "real network" would

probably be smaller if WRF assumed "cropland/grassland" for the grid-cells that represent the 411 sites. Since the WRF-simulations are all performed with the same simplified landscape, all networks are located in the same reference landscape derived by the strategy of dominant land-cover.

The artificial networks with 200 or more sites in all their ten setups represent terrain elevation well (Fig. 5.3b). In the ten 100-sites-networks, sites are, on average, located (up to 41 m) higher than in the reference landscape. The real network significantly over-represents by about 8 % sites that represent areas with elevation between 100 and 300 m and under-represent by about 6 % and 3 % sites with elevation <100m and elevation between 300 and 500 m , respectively (Fig. 5.3b).

Except for the ten 100-sites-networks and the real network, all networks also represent the frequency of soil-type occurrence well within ±5 %. The ten 100-sites-networks over-represent clay-loam (CL) (up to 7 %). The real network significantly under-represents loam (L) and over-represents CL (Fig. 5.3c). These misrepresentations of soil-types may cause some uncertainty in regional averages of soil-temperature with consequences for other quantities (e.g. 2-m temperature) derived from the real network.

5.3.2 General findings

Overall, networks with 200 or more randomly distributed sites reproduce the reference regional averages of all quantities in all setups well, while the real network has some difficulties to capture them (Table 5.4). All ten 100-sites-networks have difficulty to capture the regional averages of SLP, but reproduce the regional averages of all other quantities well. For the real networks regional averages of 2-m temperature, relative humidity, precipitation, shortwave and long-wave downward radiation differ most from the reference regional averages during strong convective situations in July and frontal passages in December no matter which year. In July, the real network has difficulties in correctly representing convective situations (as evidenced by satellite data), while its December regional averages are temporally biased during frontal passages (e.g. Fig. 5.4).

Due to its non-random site distribution, the majority of the sites within a $2.8^\circ \times 2.8^\circ$ area can be passed by fronts within short time.

All networks with randomly distributed sites typically reproduce regional averages with lower errors (RMSEs, biases, SDEs) than the real network. While SDEs, biases, and RMSEs for these 40 networks show no distinct area of higher or lower values (therefore only shown for SLP), those of the real network do (e.g. Fig. 5.5). Regional averages from the real network have high systematic and random errors for all quantities in $2.8^\circ \times 2.8^\circ$ areas located mainly over mountains and/or land-water boundaries (therefore only shown for biases). SDEs, biases, and RMSEs between the reference regional averages and regional averages derived from the real network show similar spatial and temporal behavior in all three years (Table 5.4). In the following, when quantifying errors or skills we give the worst correlation, WIA, absolute bias, RMSE and SDE.

5.3.3 Sea-level pressure

All networks with 200 or more randomly distributed sites reproduce the regional SLP-averages and their temporal evolution well (Fig. 5.4) with biases below the typical accuracy of routine measurements, correlation-skill scores >0.905 (2007) and WIA>0.901 (2007) in all of their ten setups. For the ten 100-sites-networks the low density strongly affects capturing the phase and amplitude, while for the real network the non-random site distribution causes temporal biases (up to -4.8 hPa) because fronts pass a majority of "sites" within short time (Figs. 5.2, 5.4c-d; Table 5.4).

The ten 100-sites-networks over- and underestimate the reference regional averages up to ± 6 hPa and ± 5 hPa in July and December, respectively (Fig. 5.4). Their representing a higher elevated landscape than the reference landscape is a major reason. The real network shows extremely high SLP-biases (about ± 160 hPa in July, ± 140 hPa in December) along the coasts, over the mountains and Arkhangel'sk (Fig. 5.5c-d, Table 5.5). The lower December- than July-biases of the real network result from the lower or even negative environmental lapse rate in the former ($0.4\ \mathrm{K}/100\ \mathrm{m}$, on average) than latter month ($0.7\ \mathrm{K}/100\ \mathrm{m}$, on average) due to the excessive radiative cooling over snow

(Fig. 5.6). Similar was found by Barry and Chorley (1992), for instance, over central Canada and eastern Siberia and by Huang et al. (2008) who examined the relationships between near-surface temperature, lapse rate and solar radiation. The spatial variation in environmental lapse rate is smaller in December due to the lower horizontal heterogeneity of surface temperature and moisture conditions than in July. Similar was found by Rolland (2002) when investigating the seasonal and spatial variation of lapse rates in Alpine regions. Consequently, the real network's misrepresentation of terrain affects SLP-regional averages stronger in July than in December.

According to the correlation skill scores and WIA regional SLP averages derived from the real network only marginally agree with the reference regional averages (Table 5.4). Along coasts, regional SLP-averages are even negatively correlated with the reference. The fact that sites of the real network represent an, on average, about 29 m flatter terrain than the reference landscape (Fig. 5.3b) and the strong inverse correlation (>-0.985) between SLP biases and spatial differences between terrain elevation of the real network and the reference (Fig. 5.5g-h) explain the overall strong systematic errors in regional SLP-averages. Reducing pressure measured at mountain sites to SLP assumes an average temperature between the sea-level surface and the site that may be incorrect.

RMSEs and SDEs for the real network are greatest over Arkhangel'sk and Sayan Mts. in all months (Table 5.5). SLP-RMSEs are more than ten times higher than the SDEs (Table 5.4). This means random errors are relatively small and the misrepresented terrain causes most of the regional SLP-errors. RMSEs and SDEs of the real network are on average higher in December than in July (Table 5.4).

5.3.4 10-m wind-speed

The 40 networks with randomly distributed sites well reproduce the phase and amplitude of the reference regional wind-speed averages with errors smaller than the typical errors of routine measurements, correlation-skill scores >0.926 and WIA>0.957 (Fig. 5.7).

The real network underestimates the amplitude up to $0.7\,m/s$ and $0.5\,m/s$ in July and December, respectively, and generally has difficulties in reproducing the phase (Fig. 5.7). Overall correlation (WIA) exceeds 0.515 and 0.432 (0.468 and 0.481) in July and December, respectively (Table 5.4), i.e. about 55% (45%) lower than for the networks with randomly distributed sites. In both months even negative correlations occur along the coasts and over the Ural Mts. indicating that the real network strongly misrepresents the wind field conditions at boundaries of smooth to rough or vice versa and in complex terrain. In December, regional averages of wind-speed derived from the real network become biased when frontal systems pass the majority of the non-uniformly distributed sites.

For the real network, July and December wind-speed absolute biases reach up to $3.2\,m/s$ and $6\,m/s$, respectively, along the coasts and over the Central Siberian Uplands (Fig. 5.8c-d, Table 5.5). Here also RMSEs and SDEs are highest (Table 5.5). July biases and RMSEs are about half the magnitude of those in December (Table 5.4). The systematic errors may be attributed to differences in surface and terrain roughness represented by the real network and the reference network (Fig. 5.3a). Except for the coasts, these $2.8^{\circ} \times 2.8^{\circ}$ areas represent complex terrain. The non-uniform distribution of the real network has difficulties to represent the wind direction/"channeling" situation correctly and hence, wind-speed. In the $2.8^{\circ} \times 2.8^{\circ}$ areas over water, the surface is relatively smoother and wind-speed is greater than over land. Thus, the reference regional averages exceed those calculated from the real network that only considers land "sites". In general, biases exceed the SDEs indicating that systematic errors due to misrepresentation of surface roughness and terrain by the real network dominate the RMSEs (Table 5.4).

5.3.5 2-m temperatures

The 40 artificial networks well reproduce the reference regional 2-m temperatures averages with biases below the typical accuracy of measurements, correlation-skill scores

>0.896 and WIA>0.902 (Fig. 5.9). The real network has tremendous difficulties in capturing regional 2-m temperature averages (Fig. 5.9).

In July, based on the real network regional 2-m temperature averages are overestimated up to about 8 K and underestimated up to about 12 K along coasts and over mountains (Table 5.5); in December, overestimates and underestimates in these areas are twice as high than in July (Fig. 5.10c-d). In both months the RMSEs and SDEs of the real network are also highest along the coasts and over the mountains (Table 5.5). July-RMSEs and SDEs reach up to 10 K and 5.2 K respectively; they more than double for December (Table 5.4).

For the real network, errors are higher in December than in July because of the strong influence of snow-covered and snow-free surfaces on 2-m temperatures (Table 5.4). RMSEs and SDEs are comparatively higher in the early afternoon than at other times in July due to the then strong convection not being well represented by the real network. No such obvious pattern exists for RMSE and SDE in December because of the more homogenous temperature distribution under winter high and low pressure than in summer convective situations.

These findings indicate that random errors play a role, but misrepresentation of the landscape by the real network introduces great systematic errors in regional averages of 2-m temperatures for $2.8° \times 2.8°$ areas in complex elevated terrain or that include both water and land. The former finding well agrees with Peterson (2006) and Pielke et al. (2007). The real network's failure to represent terrain elevation affects the representation of temperature distribution because temperature typically decreases with height (section 5.3.3). Furthermore, the real network has about 15 % more MF sites and 10 % less "water"-sites than required to represent the reference landscape (Fig. 5.3a). During the day in July MF heats less strong than areas covered by low vegetation; surface temperatures of lakes and the ocean are typically lower than those of the adjacent vegetation. These facts partly explain the higher regional temperature averages derived from the real network in coastal and shore areas, and the lower values in mountainous forest-covered areas than those obtained from the reference network in July. The real

network's misrepresentation of terrain elevation adds to discrepancies in regional temperature averages. In December, open water is relatively warmer than adjacent snow-covered areas. Thus, heat-fluxes from open water to the atmosphere lead to warmer air than over snow-covered land. Furthermore, lakes are frozen, homogeneously snow-covered and have relatively high albedo; high vegetation sticks out of snow, for which albedo is lower than over small entirely snow-covered vegetation. Albedo, however, affects 2-m temperature via the snow-temperature-albedo feedback. Brighter surfaces reflect more incoming radiation than relatively darker surfaces. Consequently, December regional averages based on the real network, wherein MF is over-represented at the cost of totally snow-covered small or no vegetation, are higher than those of the reference network. In the case of $2.8^\circ \times 2.8^\circ$ areas located in coastal regions the fact that the water in these areas is not completely ice-covered plays a role.

The real network has difficulties in reproducing the phase especially on days with frontal passages (e.g. July 11 and December 21, 2005) and overestimates the amplitude up to 1.9 K and 1.4 K in July and December, respectively (Fig. 5.9 c-d). The systematic errors in the temporal course of 2-m temperature regional averages derived from the real network may partly be explained by misrepresentation of incoming solar radiation (section 5.3.8).

All networks with randomly distributed sites reproduce the regional averages of maximum 2-m temperatures well for the majority of the $2.8^\circ \times 2.8^\circ$ areas in both months of all three years whereas the real network has substantial difficulties in December (Table 5.4). Typically errors in the regional averages obtained from the real network are lower in July than in December for all three years. Consequently, WIA and correlation are comparatively higher in July than December, i.e. July regional averages from the real network are more reliable than those derived for December.

The 40 networks with randomly distributed sites capture the regional averages of minimum 2-m temperatures well in both months in all three years whereas the real network shows strong biases of up to about 20 K along coasts, over mountains and south

of Arkangel'sk (Table 5.5). RMSEs and SDEs are about twice as high in December than July for the reasons discussed earlier.

In summary, systematic errors due to misrepresentation of the landscape by the real network strongly contribute to RMSEs in regional averages of maximum and minimum temperatures derived from the real network.

5.3.6 Relative humidity

In both months of all three years the 40 networks with randomly distributed sites reproduce the temporal evolution of relative humidity regional averages well (not shown). Biases are below the typical accuracy of measurements. Even the lowest correlation-skill scores and WIA still exceed 0.932. While the real network also acceptably reproduces the amplitude, it has appreciable difficulties with the phase.

Like for regional SLP- and temperature averages regional relative humidity averages derived from the real network show high errors over mountainous and coastal areas (Fig. 5.11c-d; Table 5.5). Errors in the regional averages obtained from the real network are, on average, higher in July than in December (Table 5.4). Consequently, correlation and WIA are as low as 0.490 and 0.462 in July (0.566 and 0.627 in December), respectively (Table 5.4). The highest biases and RMSEs in July are about twice as high as the values found for December (Table 5.4) because of the greater spatial differences in relative humidity in the former than latter month. The lower average correlation (>0.490) in July than December (>0.566) in conjunction with the higher July biases (Table 5.4) suggest misrepresentation of the convective situation as a contributing factor. In July and December even negative correlations between the regional relative humidity averages derived from the real and reference network occur over the mountains, along the shores of the Baikal Lake and the coasts. The nearly similar SDEs in July (16 %) and December (12 %) indicate a similar contribution of random errors in both months. The, on average, higher absolute values of biases than SDEs imply that systematic errors due to misrepresentation of the landscape contribute greatly to the RMSEs of relative humidity.

The reasons for these findings are manifold. The disagreement for areas with substantial water fraction results from the differences in surface moisture and water supply to the atmosphere for water and land areas. While over water saturation deficit and wind speed mainly determine the water supply to the atmosphere, over land vegetative controls, soil moisture and soil-type also impact the water supply and hence relative humidity. Furthermore, since the exchange of heat and moisture at the vegetation-atmosphere interface depends on vegetation-type, differences in the vegetation represented cause bias in relative humidity; differences in terrain representation may strongly affect relative humidity due to temperature differences. As aforementioned the non-random site distribution in the real network misrepresents the MF and WB frequency with consequences for the exchange of heat and moisture at the surface-atmosphere interface. Consequently, the real network cannot capture the distribution of relative humidity well. As will be discussed in sections 5.3.7 and 5.3.8, this shortcoming has consequences for convection, precipitation, shortwave and long-wave radiation with feedback to 2-m temperature. Note that relative humidity non-linearly depends on temperature; at relatively low temperatures and same specific humidity a 1 K increase in temperature, for instance, causes a greater decrease in relative humidity than at relatively high temperatures. In all years examined, the environmental lapse rate close to the surface is $\sim 0.3\ \mathrm{K/100\ m}$ stronger in July than December (Fig. 5.6).

5.3.7 Precipitation

The 40 networks with randomly distributed sites capture the temporal evolution of regional precipitation averages well with biases below the typical errors of routine measurements, WIA>0.905 and correlations >0.850. While the real network also well reproduces the amplitude, it has difficulties to capture the phase, especially during the frontal passages.

For the real network the greatest errors (biases, SDEs, RMSEs) in regional precipitation averages occur in $2.8^\circ \times 2.8^\circ$ areas that represent water or complex terrain. SDEs are only slightly lower than RMSEs ($15\ \mathrm{mm/d}$ vs. $15.5\ \mathrm{mm/d}$ in July; $12\ \mathrm{mm/d}$

vs. 13 mm/d in December; cf. Table 5.4) indicating that random errors dominate the performance of the real network in reproducing regional precipitation averages. The highest biases reach up to −4.8 mm/d and 5.1 mm/d in July (Fig. 5.12c-d; Table 5.4).

On average, biases, RMSEs and SDEs of the real network are higher in July than in December for all three years (Table 5.4). The higher July than December biases mean that misrepresentation of terrain elevation has a stronger impact than that of the convective situation. The greater precipitation biases at high elevation and greater biases in summer than in winter well agree with results from Groisman and Legates (1994) who found similar behavior for US meteorological networks. Misrepresentation of terrain height yields systematic errors related to precipitation caused by orographic lifting, while along the coasts the misrepresentation of atmospheric moisture supply goes along with misrepresentation of precipitation.

On average, in July correlation-skill scores are appreciably lower than in December because the real network misrepresents convection and hence convective precipitation. Moreover, there is more precipitation in July than December. With more areas receiving no precipitation the likelihood to "correctly" obtain zero precipitation for the regional averages by pure chance increases. The notably lower WIA than correlation skill scores (Table 5.4) indicate an offset in the regional precipitation averages featured by the real and reference network.

5.3.8 Downward radiation

All 40 networks with randomly distributed sites well capture the temporal evolution of shortwave radiation regional averages. Regional averages of shortwave radiation derived from the artificial networks have smaller biases than the typical measurement errors; correlation-skill scores and WIA exceed 0.932 and 0.931, respectively.

The real network has notable difficulties in reproducing the regional averages of shortwave radiation ($r>0.315$, $WIA>0.149$; Table 5.4). It overestimates regional averages up to about $100 Wm^{-2}$ along coasts and underestimates up to about $120\ Wm^{-2}$ over mountains in July; whereas in December, overestimates reach up to $100\ Wm^{-2}$ over

Arkhangel'sk (Fig. 5.13c-d, Table 5.5). These systematic errors can be explained as follows: In July the misrepresentation of the landscape by the real network (Fig. 5.3a) leads to an inadequate representation of the regional exchange of heat and moisture at the vegetation-atmosphere interface with consequences for relative humidity and temperature (sections 5.3.6, 5.3.7) and feedback on cloud formation. Note that cloud occurrence differs slightly among the artificial networks and the reference network, but strongly deviates from the reference for the real network. Differences in cloudiness in turn affect incoming shortwave radiation. Thus, the placement of sites in the real network causes to misrepresent cloudiness especially in partly ocean-covered $2.8^{\circ} \times 2.8^{\circ}$ areas under convective situations as evidenced from satellite imaginary. Shifts in the timing of high and low insolation occur. In December the real network's misrepresentation of the reference terrain height feeds back to misrepresentation of temperature and humidity distributions with impacts for cloudiness, and finally shortwave radiation. Overall, the real network underestimates shortwave radiation regional averages up to $19\,\mathrm{Wm}^{-2}$ and $14\,\mathrm{Wm}^{-2}$ in July and December, respectively (Table 5.4).

For the real network, shortwave radiation RMSEs are high along coasts (up to $180\mathrm{Wm}^{-2}$) in July and over Arkhangel'sk (up to $100\,\mathrm{Wm}^{-2}$) in December (Table 5.5). July-SDEs of shortwave radiation are greatest (up to $150\,\mathrm{Wm}^{-2}$) along coasts and over high mountains (Table 5.5); December-SDEs are greatest (up to $26\,\mathrm{Wm}^{-2}$) over mountains (Table 5.5). On average, SDEs are about $80\,\mathrm{Wm}^{-2}$ higher in July than in December. The higher absolute values of biases than SDE values imply that systematic errors due to misrepresentation of the reference landscape and in July convection by the real network dominate RMSEs of shortwave downward radiation.

In both months of all three years all 40 artificial networks well reproduce the temporal evolution of regional long-wave radiation averages with biases below the typical accuracy of measurement errors. Thus, WIA and correlation-skill scores exceed 0.905 and 0.945, respectively. As documented by the skill scores (Table 5.4) the real

network has some difficulties to reproduce the phase and regional averages of long-wave radiation.

Regional averages of long-wave radiation derived from the real network are biased during frontal passages. Thus, due to the non-random site distribution of the real network a great majority of the sites is passed at nearly the same time shifting the averages towards lower/higher values than the reference regional averages. The real network overestimates and underestimates the reference regional averages up to ± 60 Wm^{-2} along coasts and over mountains in July; in December it overestimates (underestimates) up to 80 Wm^{-2} (100 Wm^{-2}) over mountains (along the coast of the Sea of Okhotsk) (Fig. 5.14c-d, Table 5.5). The reasons for these systematic errors in long-wave radiation are similar to those for shortwave radiation. In addition, misrepresentation of terrain height can contribute to misrepresentation of snow-cover with consequences for temperature (via the albedo-temperature effect), moisture, cloud and finally long-wave radiation distribution in December.

In July for the real network long-wave radiation RMSEs are highest (up to 68 Wm^{-2}) over Arkhangel'sk; SDEs are greatest (up to 45 Wm^{-2}) over mountains (Table 5.5). In December, for the real network RMSEs and SDEs are highest (up to 100 Wm^{-2} and 56 Wm^{-2}, respectively) over mountains and along the coasts (Table 5.5).

Errors for the real network are higher in December than in July for most of the $2.8^{\circ} \times 2.8^{\circ}$ areas. Consequently, correlation between the averages derived from the real and reference network are notably lower in December than in July; similar is true for WIA (Table 5.4).

The nearly equal SDEs and absolute biases in regional averages of long-wave radiation found for the real network indicate that systematic and random errors contribute nearly equally to RMSEs (Table 5.4).

5.3.9 Soil-temperature

Generally, the 40 networks with randomly distributed sites reproduce regional soil-temperature averages at all depths and well capture the temporal evolution in the upper soil (Fig. 5.15) with correlation skill scores >0.943 and WIA>0.921.

The real network overestimates soil-temperature amplitudes (up to 1.4 K and 1.8 K in July and December, respectively). It has notable difficulties to capture the phase. For example, on July 11 and 28, 2005 regional averages of upper soil-temperatures fail to show the cold-snap seen in the reference regional averages (Fig. 5.15). These phase differences occur when frontal systems come through and pass the majority of sites within short time due to the non-random site distribution. Generally, biases are greatest along the coast of Barents Sea and over most mountains in July and along the coasts and over mountains in December (Fig. 5.16c-d, Table 5.5). On average, in July 2005, for instance, regional soil-temperature averages are overestimated by 2.3 K, 1.5 K, 0.8 K and 1.7 K, at 0.05 m, 0.2 m, 0.4 m and 1.6 m depths, respectively; in December 2005, the real network, overestimates regional soil-temperature averages by 2.1 K, 1.8 K, 1.3 K and 1 K at these depths. Biases of regional soil-temperature averages decrease with depth for the real network because the differences related to differences in vegetation, terrain height and atmospheric conditions between the real and reference network become less important for deeper than upper soil layers. The higher bias found for upper than lower soil layers may be partly due to misrepresentation of terrain height, vegetation and atmospheric conditions by the real network. Differences in vegetation cover/fraction and terrain elevation, namely, have consequences for soil heating. At all depths, some bias stems from the misrepresentation of the soil-type distribution by the real network. The aforementioned difficulties in capturing the temporal evolution of soil temperatures in the upper soil also result from differences between the soil-heat capacity and thermal conductivity of soils represented by the real network and those of the reference landscape. As shown by Mölders et al. (2005) small differences in these parameters can significantly (at the 95% or higher confidence level) affect soil-temperatures. Soils with high sand fraction heat/cool much quicker than those with low

sand fraction. In December the high bias found for the real network is also affected by regional differences in snow-cover and/or snow-depth. Here failure to represent terrain height and vegetation distribution plays a role for snow-conditions. More MF instead of low vegetation, for instance, means a lower albedo and snow-depth with consequences for insulation of the soil. Secondary differences may be associated with the temperature-albedo feedback. As precipitation increases with height, differences in represented snow-conditions may also occur.

Consequently, in the upper soil, soil-temperatures obtained from real network are poorly correlated (>0.315) with the reference regional soil-temperature averages and WIA>0.265 in both months. In July and December RMSEs and SDEs for the real network are highest (up to 18 K, 4 K) over the mountains and along the coast of the Barents Sea (Table 5.5). In 2005, for instance, 0.05 m, 0.2 m, 0.4 m and 1.6 m depth, July-RMSEs amount 4.4 K, 3.5 K, 3.2 K, 3.9 K and July-SDEs at these depths are 2.9 K, 1.7 K, 0.7 K, and 0.2 K; December-RMSEs reach 5.8 K, 4.6 K, 3.6 K, and 2.7 K and December-SDEs are 3.1 K, 2 K, 0.8 K and 0.2 K at these depths. The higher absolute biases than SDEs (Table 5.4) suggest that systematic errors due to misrepresentation of soil-type mainly contribute to RMSEs.

In the natural landscape differences between the regional averages derived from the real network and the true regional averages may be even greater than in our theoretical study because the real network was designed for agricultural purposes, i.e. the real network represents the fertile soils within the $2.8^\circ \times 2.8^\circ$ areas. Consequently, it may be even more biased to a soil-type than in the simplified WRF-created landscape assumed in this case study.

Note that PaiMazumder et al. (2008) showed that in December biases between a GCM-simulated and gridded soil-temperature climatology reach up to 6 K at 0.2 m depth of which about 2.5 K bias may result from incorrectly simulated atmospheric forcing. Considering the results of our case study uncertainty due network design can

explain about 2 K of their total bias in winter; thus about 1.5 K of their bias may be attributed to measurement errors and/or model deficits.

5.4 Conclusions

Simulations performed with the Weather Research and Forecasting model over Russia for July and December 2005, 2006 and 2007 are used to produce a reference dataset to examine the degree of uncertainty in regional averages caused by network density and/or design. Ten networks with four different densities of randomly distributed sites (100, 200, 400, and 500) are assumed. The WRF quantity simulated for the location of an assumed "site" of these 40 networks is assumed as a "measurement" within the respective network. Regional averages valid for $2.8^\circ \times 2.8^\circ$ areas are calculated based on the values of the "sites" that fall within these $2.8^\circ \times 2.8^\circ$ areas. These regional averages are compared to the reference regional averages that are determined based on all WRF-simulated values within a $2.8^\circ \times 2.8^\circ$ area. Furthermore, regional averages obtained from WRF-simulated values at the locations of an existing network with 411 sites ("real network") are compared to the reference regional averages.

Networks with 200 or more randomly distributed sites reliably reproduce regional averages of the examined quantities with errors smaller than the typical accuracy of measurements and show high correlation values and Willmott's index of agreement. The ten 100-sites-networks have difficulties in capturing the regional averages of SLP due to their higher terrain elevation than the reference landscape.

The real network has difficulties in capturing the reference regional averages of all quantities examined. The reasons differ for the different state variables and fluxes, with sometimes secondary effects involved. Historically the real network was designed to collect soil-temperature measurements for agricultural purposes. Thus, its soil-type distribution is skewed towards more fertile soils than the soil-type distribution in the reference landscape. The differences in soil physical properties (e.g. heat capacity, conductivity) lead to systematic error in regional soil-temperature averages determined from the real network with biases up to about 20 K . The non-random site distribution of

the real network also yields temporal offsets in soil-temperature, SLP, precipitation, shortwave and long-wave radiation during frontal passages when the majority of sites within a $2.8° \times 2.8°$ area are passed nearly at once. Since the exchange of heat and moisture at the vegetation-atmosphere interface affects 2-m temperature and relative humidity, misrepresentation of vegetation frequency, soil-type and terrain elevation propagate into misrepresentation of convection, precipitation, shortwave and long-wave radiation. Convective activity over forest and cropland, for instance, strongly differ, for which the non-randomly distributed real network cannot well represent the convective situation of $2.8° \times 2.8°$ areas. The results also show that for most quantities there are geographic trends in regional averages determined from the real network. Errors are lower for regional averages over flatland than low mountain ranges that again are lower than for high mountain ranges. Furthermore, errors in regional averages are greater in coastal areas than in areas along the shores of large lakes. No such geographical trends exist for random distributed networks. Thus, one may conclude that high priority should be given to random placement of sites when designing new networks if possible.

Our case study shows that non-random network design like low site density can introduce substantial uncertainty in gridded data and that networks with randomly distributed sites might only need about half the points of the non-random distributed real network over Russia to determine gridded data. However, maintenance of such networks with randomly distributed sites can be extremely expensive, especially when sampling is to be performed over several decades, because many of the sites would be difficult to access in remote areas.

Based on these findings we further conclude that when evaluating GCMs with gridded data from "imperfect" existing networks or networks that where not designed with this purpose in mind one has to develop intelligent strategies to guarantee meaningful conclusions on model performance and for model improvement. Similarly in determining regional averages from non-random networks strategies have to be developed to assess and remove geographical/temporal trends if possible. Sampling from model generated values as demonstrated in our study can help to evaluate geographical

and temporal trends. However, additional facts may have to be considered. In the case of using soil-temperature from the real network for evaluation, for instance, one could restrict the comparison to the patches within a GCM grid-cell that represent agriculturally used land.

Acknowledgements

We thank U.S. Bhatt, P.A. Bieniek, M.E. Brown, T. Fathauer, G. Kramm, S.E. Porter, J.E. Walsh and the anonymous reviewers for fruitful discussion, ARSC and NCAR for computational support. This research was supported by EPSCoR-grant 0701898, NSF cooperative agreements OPP-0327664 and ARC0652838.

References

Anthes, R.A., 1984: Enhancement of convectionale precipitation by mesoscale variations in vegetative covering in semiarid regions. *J. Clim. Appl. Met.*, **23**, 541-554.

Anthes, R.A., Y.H. Kuo, E.Y. Hsie, S. Low-Nam, and T.W. Bettge, 1989: Estimation of skill and uncertainty in regional numerical models. *Quart. J. Roy. Meteorol. Soc.*, **111**, 763-806.

Avissar, R., and R.A. Pielke, 1989: A parameterization of heterogeneous land surface for atmospheric numerical models and its impact on regional meteorology. *Mon. Wea. Rev.*, **117**, 2113-2136.

Baker, P.H., G.H. Galbraith, R.C. McLean, C. Hunter, and C.H. Sanders, 2006: The measurement and prediction of conditions with in indoor microenvironments. *Indoor. Built. Environ.*, **15**, 357-564.

Barry, R.G., and R.J. Chorley, 1992: Atmosphere, Weather and Climate (6^{th} ed.). London, UK: Routledge, 392pp.

Bauer, M., A.D. Del Genio, and J.R. Lanzante, 2002: Observed and simulated temperature-humidity relationship: Sensitivity to sampling analysis. *J. Climate*, **15**, 203-215.

Cannon, A.J. and P.H. Whitfield, 2002: Downscaling recent streamflow conditions in British Columbia, Canada using ensemble neural network models. *J. Hydrol.*, **259**, 136-151.

Cess, R.D., E.G. Dutton, J.J. DeLuisi, and F. Jiang, 1991: Determining surface solar absorption from broadband satellite measurements for clear skies: Comparison with surface measurement. *J. Climate*, **4**, 236-247.

Chase, T.N., R.A. Pielke, T.G.F. Kittel, R. Nemani, and S.W. Running, 1996: The sensitivity of a general circulation model to global changes in leaf area index. *J. Geophys. Res.*, **101**, 7393-7408.

Chou, M.D., and M.J. Suarez, 1994: An efficient thermal infrared radiation parameterization for use in general circulation models. *NASA Tech. Memo.104606*, **3**, 85pp.

Court, A., and M.T. Bare, 1984: Basin precipitation estimates by Bethlahmy's two-axis method. *J. Hydrol.*, **68**, 149-158.

Creutin, J.D. and C. Obled, 1982: Objective analysis and mapping techniques for rainfall fields: an objective comparison. *Water Resour. Res.*, **15**, 1752-1762.

Daily, C., W.P. Gibson, G.H. Taylor, M.K. Doggett, and J.I. Smith, 2007: Observer bias in daily precipitation measurement at United States cooperative network stations. *Bull. Amer. Meteorol. Soc.*, **88**, 899-912.

Entekhabi, D., and K.L. Brubaker, 1995: An analytical approach to modeling land-atmosphere interaction. 2. Stochastic formulation. *Water Resource Res.*, **31**, 633-643.

Frei, C., and C. Schär, 1998: A precipitation climatology of the Alps from high resolution rain-gauge observation. *Int. J. Climatol.*, **18**, 873-900.

Goody, R., J. Anderson, T. Karl, R.B. Miller, G. North, J. Simpson, G. Stephens, and W. Washington, 2002: Why monitor the climate? *Bull. Amer. Meteorol. Soc.*, **83**, 873-878.

Grell, G.A., and D. Devenyi, 2002: A generalized approach to parameterizing convection combining ensemble and data assimilation techniques. *Geophys. Res. Lett.*, **29**, 1693.

Groisman, P.Y., V.V. Koknaeva, T.A. Belokrylova, and T.R. Karl, 1991: Overcoming biases of precipitation measurement: A history of the USSR Experience. *Bull. Amer. Meteorol. Soc.*, **72,** 1725-1733.

Groisman, P.Y., and D.R. Legates, 1994: The accuracy of United States precipitation data. *Bull. Amer. Meteorol. Soc.*, **75**, 215-227.

Hanna, S.R., 1994: Mesoscale meteorological model evaluation techniques with emphasis on needs of air quality models. In: Pielke, R.A., and R.P. Pearce (Eds.), *Mesoscale modeling of the Atmosphere*, Meteorological Monographs, Boston, **25**, 47-58.

Holdaway, M.R., 1996: Spatial modeling and interpolation of monthly temperature using kriging. *Clim. Res.*, **6**, 215-225.

Huang, S., P.M. Rich, R.L. Crabtree, C.S. Potter, and P. Fu, 2008: Modeling monthly near-surface air temperature from solar radiation and lapse rate: Application over complex terrain in Yellowstone national park. *Phys. Geography*, **29**, 158-178.

Janjić, Z.I., 1996: The surface layer in the NCEP Eta Model. *Eleventh Conference on Numerical Weather Prediction*, Norfolk, VA, 19-23 August; Amer. Meteor. Soc., Boston, MA, 354-355.

Janjić, Z.I., 2002: Nonsingular implementation of the Mellor-Yamada level 2.5 scheme in the NCEP Meso model. *NCEP Office Note*, **437**, 61 pp.

Jeffery, S.J., J.O. Carter, K.B. Moodie, and A.R. Beswick, 2001: Using spatial interpolation to construct a comprehensive archive of Australian climate data. *Environ. Model. Softw.*, **16**, 309-330.

Jones, P.D., S.C.B. Raper, R.S. Bradley, H.F. Diaz, P.M. Kelly, and T.M.L. Wigley, 1986: Northern hemisphere surface temperature variation: 1851-1984. *J. Climate Appl. Meteor.*, **25**, 161-179.

Kramm, G., N. Beier, T. Foken, H. Müller, P. Schröder, and W. Seiler, 1996: A SVAT scheme for NO, NO_2, and O_3 - model description. *Meteorol. Atmos. Phys.*, **61**, 89-106.

Lebel, T., G. Bastin, C. Obled, and J.D. Creutin, 1987: On the accuracy of areal rainfall estimation: a case study. *Water Resour. Res.*, **23**, 2123-2134.

Lefohn, A.S., H.P. Knudsen, J.L. Logan, J. Simpson, and C. Bhumralkar, 1987: An evaluation of the kridging method to predict 7-h seasonal mean ozone concentration for estimating crop losses. *J. Air Pollution Contr. Asso.*, **37**, 595-602.

Li, Z., H.G. Leighton, and R.D. Cess, 1993: Surface net solar radiation estimated from satellite measurements: Comparison with tower observation. *J. Climate*, **6**, 1764-1772.

Li, Z., U.S. Bhatt, and N. Mölders, 2008: Impact of doubled CO_2 on the interaction between the regional and global water cycle in four study regions. *Clim. Dyn.*, **30**, 255-275.

Lindley, S.J., and T. Walsh, 2004: Inter-comparison of interpolated background nitrogen dioxide concentrations across Greater Manchester, UK. *Atmos. Environ.*, **39**, 2709-2724.

Luo, W., M.C. Taylor, and S.R. Parker, 2008: A comparison of spatial interpolation methods to estimate continuous wind-speed surfaces using irregularly distributed data from England and Wales. *Int. J. Climatol.*, **28**, 947-959.

MacWhorter, M.A., and R.A. Weller, 1991: Error in measurements of incoming shortwave radiation made from ships and buoys. *J. Atmos. Ocea. Technol.*, **8**, 108-117.

Mesinger, F., G. DiMego, E. Kalnay, K. Mitchell, P.C. Shafran, W. Ebisuzaki, D. Jovic, J. Woollen, E. Rogers, E.H. Berbery, M.B. Ek, Y. Fan, R. Grumbine, W. Higgins, H. Li, Y. Lin, G. Manikin, D. Parrish, and W. Shi, 2006: North American regional reanalysis . *Bull. Amer. Meteorol. Soc.*, **87**, 1073-1080.

Mitchell, R.M., and D.M. O'Brien, 1987: Error estimates for passive satellite measurement of surface pressure using absorption in the band of oxygen. *J. Atmos. Sci.*, **44**, 15, 1981-1990.

Mitchell, T.D., T.R. Carter, P.D. Jones, M. Hulme, and M. New, 2004: *A comprehensive set of high-resolution grids of the monthly climate for Europe and the globe: the observed record (1901-2000) and 16 scenarios (2002-2100).* Working paper 55, Tyndall Centre for Climate Change Research, Norwich, pp.

Mlawer, E.J., S.J. Taubman, P.D. Brown, M.J. Iacono, and S.A. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res.*, **102D**, 16663-16682.

Mölders, N., U. Haferkorn, J. Döring, and G. Kramm, 2003: Long-term numerical investigations on the water budget quantities predicted by the hydro-thermodynamic soil vegetation scheme (HTSVS) – Part II: Evaluation, sensitivity, and uncertainty. *Meteorol. Atmos. Phys.*, **84**, 137-156.

Mölders, N., M. Jankov, G. Kramm, 2005: Application of Gaussian error propagation principles for theoretical assessment of model uncertainty in simulated soil processes caused by thermal and hydraulic parameters. *J. Hydrometeorol.*, **6**, 1045-1062.

Mölders, N., and V.E. Romanovsky, 2006: Long-term evaluation of the Hydro-Thermodynamic Soil-Vegetation Scheme's frozen ground/permafrost component using observations at Barrow, Alaska. *J. Geophys. Res.*, **111**, D04105, doi:10.1029/2005JD005957.

Mölders, N., and G. Kramm, 2007: Influence of wildfire induced land-cover changes on clouds and precipitation in Interior Alaska - A case study. *Atmos. Res.*, **84**, 142-168.

Mölders, N., H. Luijting, and K. Sassen, 2008: Use of Atmospheric Radiation Measurement program data from Barrow, Alaska, for evaluation and development of snow albedo parameterizations. *Meteor. Atmos. Phys.*, **99**, 199-219.

Monin, A.S., and A.M. Obukhov, 1954: Basic laws of turbulent mixing in the surface layer of the atmosphere. *Contrib. Geophys. Inst. Acad. Sci.*, USSR, **151**, 163-187.

Ninyerola, M., X. Pons, and J.M. Roure, 2000: A methodological approach of climatological modeling of air temperature and precipitation through GIS techniques. *Int. J. Climatol.*, **20**, 1823-1841.

PaiMazumder, D., J. Miller, Z. Li, J. E. Walsh, A. Etringer, J. McCreight, T. Zhang, and N. Mölders, 2008: Evaluation of Community Climate System Model soil-temperatures using observations from Russia. *Theor. Appl. Climatol.*, **94**, 187-213.

Palutikof, J.P., J.A. Winkler, C.M. Goodess, and J.A. Andresen, 1997: The simulation of daily temperature time series from GCM output. Part I: Comparison of model data with observation. *J. Climate*, **10,** 2497-2531.

Peterson, T.C., 2006: Examination of potential biases in air temperature caused by poor station locations. *Bull. Amer. Meteorol. Soc.*, **87**, 1073-1080.

Phillips, D.L., E.H. Lee, A.A. Herstrom, W.E. Hogsett, and D.T. Tingley, 1997: Use of auxiliary data for the spatial interpolation of ozone exposure on southeastern forests. *Environmetrics*, **8**, 43-61.

Pielke Sr., R.A., J. Nielsen-Gammon, C. Davey, J. Angel, O. Bliss, M. Cai, N. Doesken, S. Fall, D. Niyogi, K. Gallo, R. Hale, K.G. Hubbard, X. Lin, H. Li, and S. Raman, 2007: Documentation of uncertainties and biases associated with surface temperature measurement sites for climate change assessment. *Bull. Amer. Meteorol. Soc.*, **88**, 913-928.

Robeson, S.M., and J.A. Doty, 2005: Identifying rogue air temperature station using cluster analysis of percentile trends. *J. Climate*, **18**, 1275-1287.

Rolland, C., 2002: Spatial and seasonal variation of air temperature lapse rates in Alpine regions. *J. Climate*, **16**, 1031-1046.

Romanovsky, V.E., T.S. Sazonova, V.T. Balobaev, N.I. Shender, and D.O. Sergueev, 2007: Past and recent changes in air and permafrost temperatures in eastern Siberia. *Global and Planetary Change*, **56**, 399-413.

Shaw, E.M., and P.P. Lynn, 1972: A real rainfall evaluation using two surface fitting techniques. *Hydrol. Sci. Bull.*, **17**, 419-433.

Skamarock, W.C., J.B. Klemp, J. Dudhia, D.O. Gill, D.M. Baker, W. Wang, and J.G. Powers, 2005: A description of the advanced research WRF version 2. *NCAR Technical Note, NCAR/TN-468+STR*, pp. 88.

Smirnova, T.G., J.M. Brown, and S.G. Benjamin, 1997: Performance of different soil model configurations in simulating ground surface temperature and surface fluxes. *Mon. Wea. Rev.*, **125**, 1870-1884.

Smirnova, T.G., J.M. Brown, S.G. Benjamin, and D. Kim, 2000: Parameterization of cold season processes in the MAPS land-surface scheme. *J. Geophys. Res.*, **105D**, 4077-4086.

Spindler, G., N. Mölders, J. Hansz, N. Beier, and G. Kramm, 1996: Determining the dry deposition of SO_2, O_3, NO, and NO_2 at the SANA core station Melpitz. *Meteorol. Zeitsch.*, **5**, 205-220.

St.-Hilaire, A., Taha B.M. J. Ouarda, M. Lachance, B. Bobée, J. Gaudet, and C. Gignac, 2003: Assessment of the impact of meteorological network density on the estimation of basin precipitation and runoff: a case study. *Hydrol. Process.*, **17**, 3561-3580.

Tabios III, G.Q., and J.D. Salas, 1985: A comparative analysis of the techniques for spatial interpolation of precipitation. *J. Amer. Water Res. Assoc.*, **21**, 365-380.

Thompson, G., R.M. Rasmussen, and K. Manning, 2004: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part I: Description and sensitivity analysis. *Mon. Wea. Rev.*, **132**, 519-542.

Tsintikidis, D., K.P. Georgakakos, J.A. Sperfslage, D.E. Smith, and T.M. Carpenter, 2002: Precipitation uncertainty and rain gauge network design within Folsom Lake watershed. *J. Hydrologic Engrg.*, **7**, 175-184.

Wilks, D.S., 1995: *Statistical methods in atmospheric sciences*. Academic Press, San Diego, 467pp.

Willmott, C.J, 1984: On the evaluation of model performance in physical geography. In: G.L. Gaile and C.J. Willmott (eds). *Spatial Statistics and Models.* Dordrecht, Holland: D. Reidel, 443-460.

Zhang, T., R.G. Barry, and D. Gilichinsky, 2001: Russian historical soil-temperature data, Boulder, Colorado, USA: National Snow and Ice Data Center, Digital media, http://nsidc.org/arcss078.html/

Zhong, S., H.J. In, X. Bian, J. Charney, W. Heilman, and B. Potter, 2005: Evaluation of real-time high-resolution MM5 predictions over the Great Lakes region. *Mon. Wea. Rev.*, **20**, 63-81.

Table 5.1 Physical packages used in the WRF-simulations to create the reference dataset assumed as an ideal reference network.

Processes	Parameterization	Reference
resolvable scale cloud and precipitation formation	mixed-phase six water class (water vapor, cloud-water, rainwater, ice, snow, graupel) bulk-microphysics parameterization	Thompson et al. (2004)
subgrid-scale convection	ensemble parameterization	Grell-Devenyi (2002)
Shortwave radiation	Goddard-scheme	Chou and Suarez (1994)
Long-wave radiation	Rapid Radiative Transfer Model	Mlawer et al. (1997)
atmospheric boundary layer	Yonsei University scheme	Skamarock et al. (2005)
Surface layer	Monin-Obukhov similarity theory	Monin and Obukhov (1954), Janjić (1996, 2002)
Soil heat and moisture transfer, frozen ground, snow temperature and moisture, transpiration, exchange of heat and moisture at the land-atmosphere interface	modified version of the Rapid Update Cycle land-surface model	Smirnova et al. (1997, 2000)

Table 5.2 Typical accuracy of routine measurements of SLP, 10m-wind-speed, 2-m temperature, relative humidity, precipitation, shortwave and long-wave radiation and soil-temperature. Note that routine measurements have greater errors than measurements of special field campaigns (cf. Spindler et al. 1996).

Quantities	Accuracy	Reference
Sea level pressure	± 2 hPa	Mitchell and O'Brien (1987)
10m-wind-speed	± 0.17 m/s	Mölders et al. (2008)
2-m temperature	± 3 K	Jones et al. (1986), Li et al. (2008)
2-m relative humidity	± 2 %	Baker et al. (2006)
Daily accumulated precipitation	± 1 mm/d	Frei and Schär (1998)
Shortwave downward radiation	± 10 Wm^{-2}	MacWhorter and Weller (1991)
Long-wave downward radiation	± 7 Wm^{-2}	Cess et al. (1991), Li et al. (1993)
Soil temperature	± 1 K	Mölders and Romanovsky (2006)

Table 5.3 Equations to calculate the performance measures (e.g. Anthes 1983, Anthes et al. 1989, Hanna 1994, Wilks 1995) used in this study. Here ϕ_i (x_i - y_i) is the difference between the regional average of a quantity obtained from the various network and "reference"-network at the i^{th} hour for a given $2.8^\circ \times 2.8^\circ$ area and n is the total number of hours within a month. Furthermore, x_i is the regional average of a quantity for a $2.8^\circ \times 2.8^\circ$ area obtained from a network and y_i is the regional average of a quantity obtained for the $2.8^\circ \times 2.8^\circ$ area from the "reference"-network for the i^{th} hour.

Skill-score	Equation
Bias	$\bar{\phi} = \frac{1}{n}\sum_{i=1}^{n}\phi_i = \frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)$
Root mean square error	$RMSE = \left(\frac{1}{n-1}\sum_{i=1}^{n}(\phi_i)^2\right)^{1/2}$
Standard deviation of error	$SDE = \left(\frac{1}{n-1}\sum_{i=1}^{n}(\phi_i - \bar{\phi})^2\right)^{1/2}$
Correlation-skill score	$r = \frac{\sum_{i=1}^{n} x_i y_i - \frac{\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n}}{\sqrt{\left(\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}\right)\left(\sum_{i=1}^{n} y_i^2 - \frac{\left(\sum_{i=1}^{n} y_i\right)^2}{n}\right)}}$
Willmott's index of agreement	$WIA = 1 - \frac{\sum_{i=1}^{n}(y_i - x_i)^2}{\sum_{i=1}^{n}\left(\left\lvert x_i - \frac{\sum_{i=1}^{n} y_i}{n}\right\rvert + \left\lvert y_i - \frac{\sum_{i=1}^{n} y_i}{n}\right\rvert\right)^2}$

Table 5.4 Range of biases, SDEs, RMSEs, correlation-skill scores and Willmott's index of agreement between regional averages for $2.8° \times 2.8°$ areas of sea level pressure (SLP), 10-m wind speed (v), 2-m temperature (T), minimum (T_{min}) and maximum temperature (T_{max}), relative humidity (RH), precipitation (P), shortwave radiation ($R_s^{\downarrow}$), long-wave radiation ($R_l^{\downarrow}$), and soil-temperature at 0.2 m depth (T_s) as obtained from the real network and the "reference"-network for July and December of all three years. The first, second and third value with in the bracket represent mean value of a given performance skill for 2005, 2006 and 2007, respectively.

	Bias		SDE		RMSE		Correlation-skill score		WIA	
	July	December	July	December	July	December	July	December	July	December
SLP (hPa)	-160.3 to 159.8 (-4.2, -4.8, -3.7)	-140.0 to 139.7 (-4.0, -4.8, -4.5)	1.0 to 10.1 (4.2, 3.8, 4.9)	1.9 to 19.8 (7.9, 8.6, 9.0)	20.1 to 259.8 (31.1, 29.5, 35.0)	20.0 to 280.3 (33.7, 31.5, 30.2)	-0.609 to 0.817 (0.375, 0.432, 0.333)	-0.551 to 0.827 (0.522, 0.525, 0.486)	0.134 to 0.473 (0.279, 0.315, 0.215)	0.123 to 0.439 (0.256, 0.301, 0.213)

Table 5.4 continued

v (m/s)	-2.8 to 3.2 (-0.7, -0.7, -0.4)	-6.0 to 5.8 (-0.5, -0.2, -0.5)	0.5 to 3.2 (1.2, 1.4, 1.3)	0.5 to 3.8 (1.5, 1.7, 2.1)	0.5 to 4.2 (1.4, 1.8, 1.5)	1.0 to 8.8 (2.0, 2.1, 2.4)	-0.569 to 0.925 (0.632, 0.515, 0.701)	-0.533 to 0.807 (0.569, 0.432, 0.607)	0.238 to 0.734 (0.565, 0.468, 0.625)	0.269 to 0.729 (0.532, 0.481, 0.595)
T (K)	-12.0 to 7.8 (1.9, 1.9, 1.5)	-24.1 to 15.9 (1.1, 1.3, 1.4)	0.8 to 5.2 (4.3, 3.9, 4.6)	1.9 to 11.1 (2.0, 2.0, 1.9)	2.0 to 9.7 (4.9, 4.7, 4.6)	1.8 to 23.8 (8.0, 7.5, 7.2)	0.419 to 0.875 (0.675, 0.515, 0.715)	0.554 to 0.770 (0.505, 0.565, 0.495)	0.367 to 0.775 (0.583, 0.467, 0.623)	0.395 to 0.801 (0.583, 0.642, 0.502)

Table 5.4 continued

T_{max} (K)	-12.0 to 6.1 (-0.3, -0.9, -1.1)	-20.3 to 21.1 (-1.1, 1.6, 1.5)	1.2 to 9.0 (4.0, 3.2, 4.1)	2.2 to 11.7 (5.1, 6,7, 4.2)	1.0 to 15.2 (5.0, 5.1, 5.8)	0 to 24 (8.0, 7.3, 8.5)	0.355 to 0.801 (0.762, 0.715, 0.648)	0.245 to 0.793 (0.574, 0.639, 0.448)	0.505 to 0.870 (0.593, 0.683, 0.515)	0.218 to 0.808 (0.448, 0.541, 0.358)
Tmin (K)	-20.1 to 6.2 (0.7, -2.1, -1.9)	-20.1 to 19.7 (-1.2, -2.7, 3.0)	0.4 to 6.8 (2.8, 2.1, 1.7)	2 to 13 (5.4, 6.2, 5.1)	1.0 to 14.1 (3.6, 2.7, 3.1)	1.1 to 23.8 (8.7, 7.1, 9.6)	0.441 to 0.987 (0.720, 0.556, 0.451)	0.546 to 0.931 (0.563, 0.673, 0.701)	0.145 to 0.861 (0.658, 0.557, 0.432)	0.307 to 0.891 (0.525, 0.673, 0.718)

Table 5.4 continued

RH (%)	-30 to +30 (-2.5, -2.3, -2.0	-16 to +16 (0.3, 0.1, 0.1)	3 to 16 (8.8, 6.2, 7.1)	2 to 12 (5.2, 4.0, 3.9)	4 to 44 (11.8, 8.6, 9.2)	2 to 20 (6.6, 4.6, 5.3)	-0.643 to 0.919 (0.490, 0.558, 0.511)	-0.496 to 0.875 (0.566, 0.594, 0.607)	0.347 to 0.856 (0.462, 0.493, 0.595)	0.393 to 0.801 (0.627, 0.715, 0.703)
P (mm/d)	-4.8 to 5.1 (-0.2, -0.2, -0.1)	-5.3 to 3.4 (-0.3, -0.5, -0.1)	1.8 to 15.2 (4.9, 4.1, 5.2)	2.2 to 11.7 (1.8, 1.5, 1.6)	1.7 to 15.5 (5.2, 5.0, 5.1)	1.0 to 12.8 (2.0, 1.6, 2.0)	-0.380 to 0.892 (0.562, 0.496, 0.521)	-0.378 to 0.954 (0.726, 0.705, 0.627)	0.269 to 0.815 (0.401, 0.364, 0.652)	0.290 to 0.891 (0.645, 0.693, 0.592)

Table 5.4 continued

$R_s^{\downarrow}$ (Wm^{-2})	-120 to 100 (19, 18, 16)	-30 to 100 (13, 11, 13)	20 to 150 (78, 81, 76)	2 to 26 (5, 5, 4)	20 to 180 (87, 91, 83)	10 to 100 (6, 8, 8)	-0.606 to 0.751 (0.149, 0.394, 0.267)	-0.337 to 0.776 (0.405, 0.569, 0.458)	0.157 to 0.697 (0.205, 0.415, 0.365)	0.202 to 0.815 (0.279, 0.627, 0.525)
$R_l^{\downarrow}$ (Wm^{-2})	-60 to 60 (5, 6, 8)	-100 to 80 (1, 2, 2)	10 to 45 (24, 30, 21)	12 to 56 (33, 38, 31)	12 to 68 (29, 35, 28)	20 to 100 (44, 46, 41)	-0.015 to 0.971 (0.882, 0677, 0.758)	-0.199 to 0.951 (0.631, 0.525, 0.615)	0.405 to 0.890 (0.734, 0.598, 0.655)	0.395 to 0.897 (0.685, 0.563, 0.607)
T_s (K)	-10.3 to 19.7 (1.5, 1.1, 1.3)	-17.7 to 11.8 (1.8, 1.1, 1.7)	1.3 to 3.9 (1.7, 1.5, 1.2)	1.1 to 4.1 (1.9, 1.1, 1.5)	1.8 to 17.6 (3.5, 3.2, 3.1)	2.2 to 18.0 (4.6, 3.9, 4.1)	0.383 to 0.681 (0.369, 0.421, 0.315)	0.253 to 0.732 (0.445, 0.523, 0.519)	0.175 to 0.703 (0.265, 0.396, 0.497)	0.167 to 0.817 (0.267, 0.425, 0.601)

Table 5.5 Locations of highest biases, RMSEs and SDEs between regional averages for $2.8^{\circ} \times 2.8^{\circ}$ areas of the various quantities obtained from the real network and the "reference"-network for July and December. Areas of highest values are independent of years.

Quantities	Bias		RMSE		SDE	
	July	December	July	December	July	December
Sea-level pressure (hPa)	Coast of the Sea of Okhotsk, Sayan Mts., Verkhoyansk Mts., Arkhangel'sk	Coasts of the Sea of Okhotsk, Sayan Mts., Altay Mts., Arkhangel'sk	Arkhangel'sk	Arkhangel'sk	Ural Mts.,Central Siberian Uplands	Central Siberian Uplands
10m-wind-speed (m/s)	Coasts of the Sea of Okhotsk, Central Siberian Uplands	Coasts of the Sea of Okhotsk, Barents Sea, Kara Sea, Irkutsk	Central Siberian Uplands, coasts of the Sea of Okhotsk	Coasts of the Sea of Okhotsk, Central Siberian Uplands	Coasts of the Sea of Okhotsk, Barents Sea	Central Siberian Uplands, coasts of the Sea of Okhotsk

Table 5.5 continued

2-m temperature (K), maximum and minimum temperatures	Coasts of the Sea of Okhotsk, Barents Sea, Kara Sea, Sayan Mts., Altay Mts.	Coasts of Sea of Okhotsk, Cherski Mts., Verkhoyansk Mts., Central Siberian Uplands	Sayan Mts., coasts of Sea of Okhotsk, Barents Sea	Coasts of Sea of Okhotsk, Ural Mts., Central Siberian Uplands	Coasts of Barents Sea, Kara Sea, Cherski Mts.	Coasts of Sea of Okhotsk, Ural Mts., Central Siberian Uplands, Irkutsk
Relative humidity (%)	Coast of Barents Sea, Sayan Mts., Altay Mts., Central Siberian Uplands	Coast of Barents Sea, Verkhoyansk Mts. Cherski Mts., Arkhangel'sk	Coast of Barents Sea, Sayan Mts., Altay Mts.	Coast of Barents Sea, Baltic Sea, Arkhangel'sk	Verkhoyansk Mts.,Ural Mts., Sayan Mts.	Sayan Mts., Altay Mts.

Table 5.5 continued

Precipitation (mm/d)	Coasts of the Sea of Japan, Barents Sea, shore of Lake Baikal, Irkutsk, Sayan Mts.	Altay Mts., Coasts of the Sea of Okhotsk, Barents Sea, Japan Sea, Central Siberian Uplands	Coasts of the Sea of Okhotsk, Sea of Japan, Irkutsk	Irkutsk, north of Altay Mts.	Coasts of Sea of Okhotsk, Sea of Japan, Irkutsk	Irkutsk, north of Altay Mts.
Shortwave radiation (Wm-2)	Coasts of Sea of Okhotsk, Barents Sea Sea of Japan, Altay Mts.	Ural Mts., Arkhangel'sk	Coasts of Sea of Okhotsk	Coasts of Barents Sea, Arkhangel'sk	Coasts of Sea of Okhotsk, Verkhoyansk Mts.	Cherski Mts., Altay Mts., Tomsk

Table 5.5 continued

Long-wave radiation (Wm^{-2})	Coasts of Barents Sea, Kara Sea, Arkhangel'sk, Altay Mts., Sayan Mts.	Coasts of the Sea of Okhotsk, Cherski Mts., Verkhoyansk Mts., Central Siberian Uplands	Arkhangel'sk	Coasts of Sea of Okhotsk, Verkhoyansk Mts., Central Siberian Uplands	Cherski Mts., Central Siberian Uplands	Coasts of Sea of Okhotsk, Sayan Mts., Arkhangel'sk
Soil-temperature (0.2 m) (K)	Coasts of Barents Sea, Sayan Mts., Ural Mts., Altay Mts.	Coasts of Sea of Okhotsk, Barents Sea, Cherski Mts., Ural Mts., Central Siberian Uplands, Tomsk	Sayan Mts., Altay Mts.	Cherski Mts., Central Siberian Uplands	Central Siberian Uplands, Coasts of Barents Sea, Ural Mts., Verkhoyansk Mts.	Central Siberian Uplands, Coasts of Barents Sea, Ural Mts.

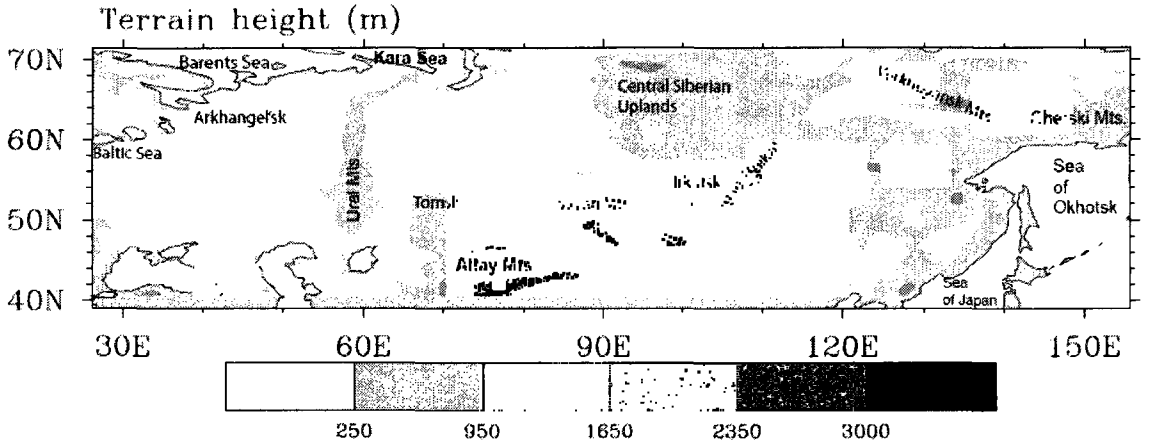


Fig.5.1 Schematic view of terrain elevation over Russia. Names refer to locations mentioned in the text.

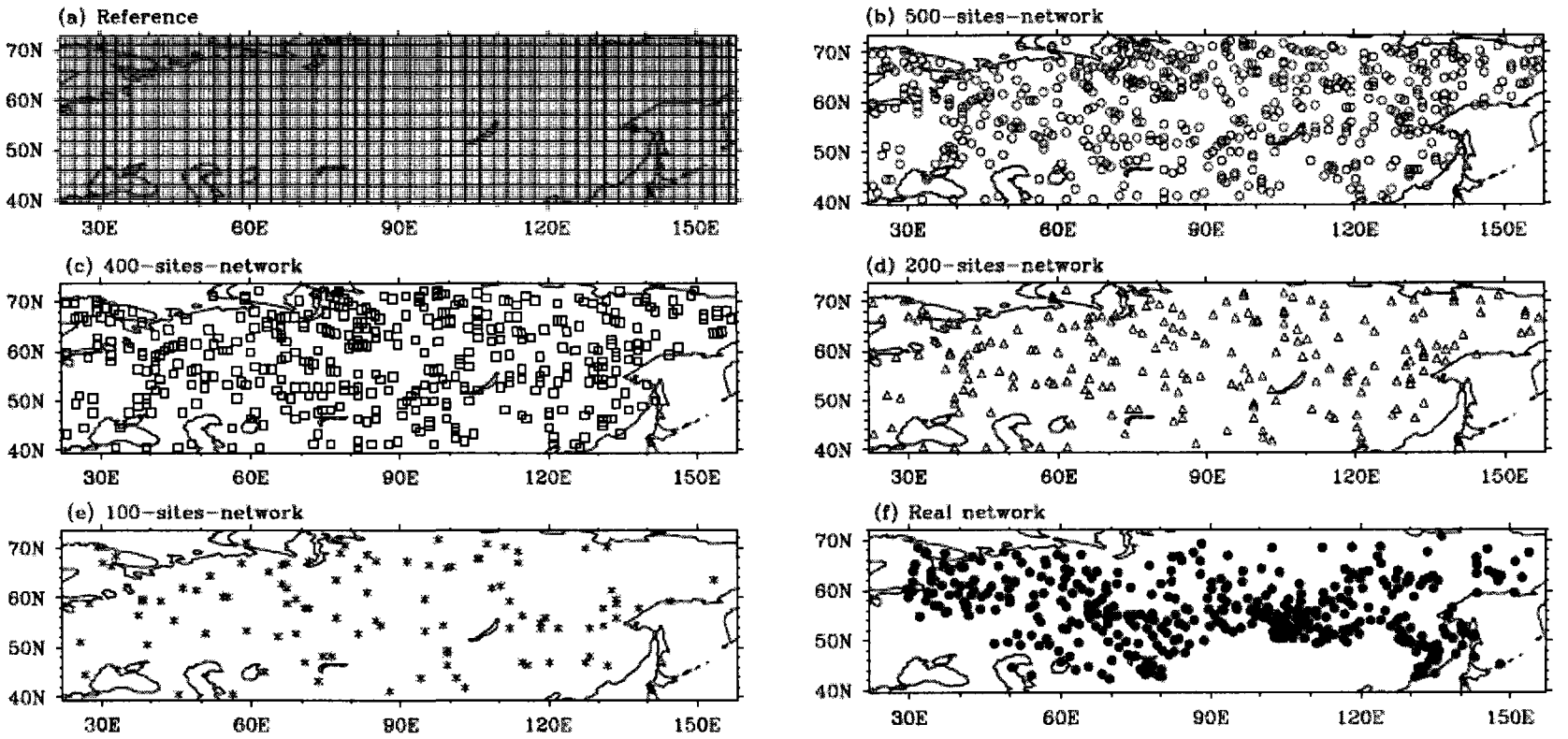


Fig. 5.2 (a) Schematic view of the $2.8° \times 2.8°$ areas and locations of "sites" in the "reference" network over Russia. Parts (b) to (e) show the location of sites for one of the ten 500- (b), 400- (c), 200- (d), and 100-sites-networks (e). (f) Locations of sites in the "real network".

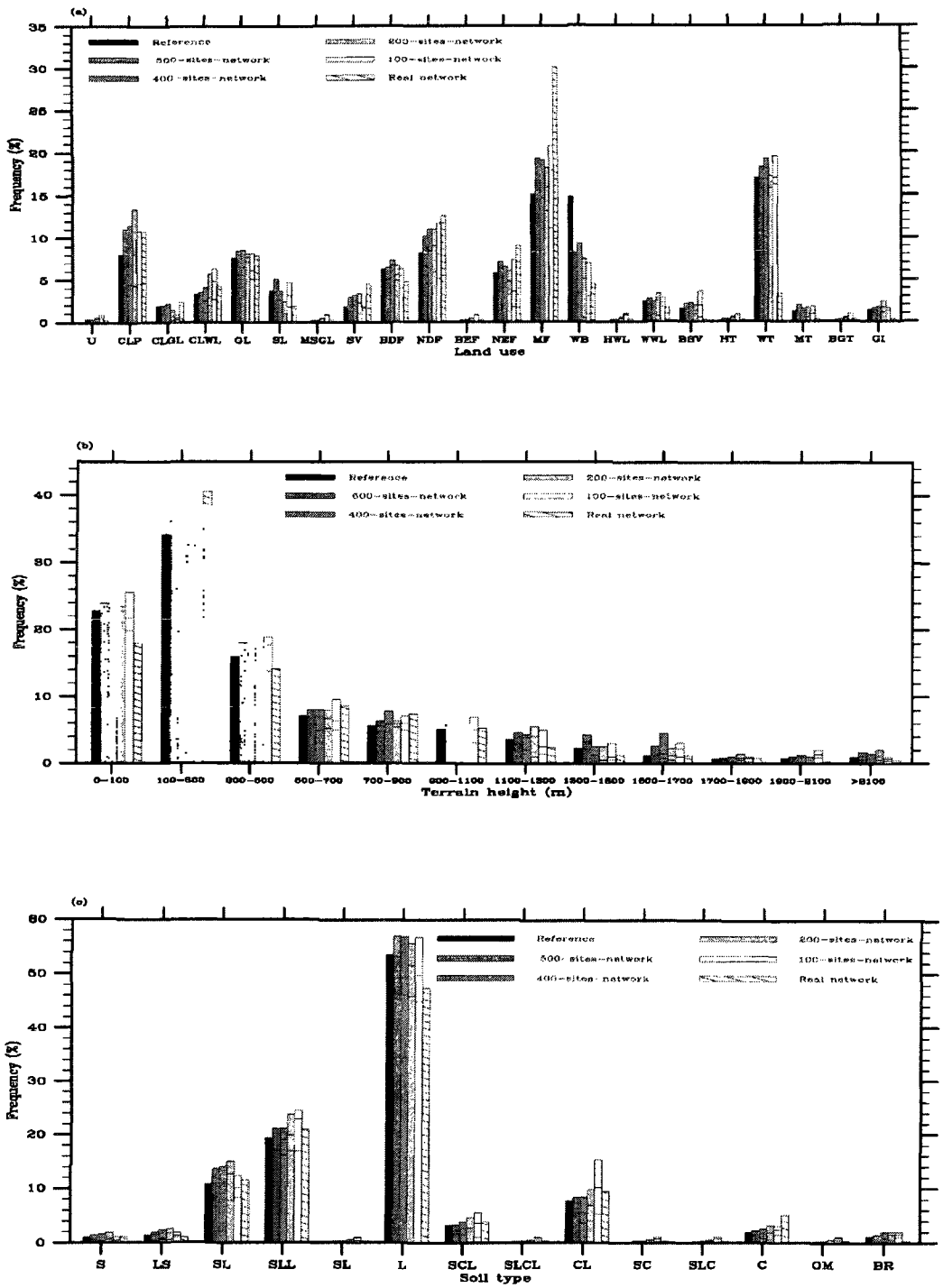


Fig. 5.3 Frequency distribution of (a) land-cover type, (b) terrain elevation and (c) soil-type or other surface type in the "reference", the 500-, 400-, 200-, 100-sites networks and real network over Russia. In the case of the artificial networks, the uppermost, mean and lowermost frequency of the ten networks are illustrated for each network density by

horizontal lines. In (a) the X-axis represents urban (U), cropland/pasture (CLP), cropland/grassland (CLGL), cropland/woodland (CLWL), grassland (GL), shrub-land (SL), mixed shrub-land/grassland (MSGL), savanna (SV), broadleaf deciduous forest (BDF), needle-leaf deciduous forest (NDF), broadleaf evergreen forest (BEF), needle-leaf evergreen forest (NEF), mixed forest (MF), water bodies (WB), herbaceous wetlands (HWL), wooded wetlands (WWL), bare or sparsely vegetated (BSV), herbaceous tundra (HT), wooded tundra (WT), mixed tundra (MT), bare ground tundra (BGT) and glacier/ice (GI). In (c), the X-axis represents sand (S), loamy sand (LS), sandy loam (SL), silt loam (SLL), silt (SL), loam (L), sandy clay loam (SCL), silty clay loam (SLCL), clay loam (CL), sandy clay (SC), silty clay (SLC), clay (C), organic material (OM) and bedrock (BR). Note that over Russia the average terrain elevation of the 500-, 400-, 200-, 100-sites-networks ranges between 405 and 421 m (with on average over all ten networks 412 m), 390 and 430 m (410 m), 389 and 454 m (415 m), and 379 and 457 m (421 m), respectively; average terrain height of the real network amounts 387 m; the average terrain height of the reference network is 416 m.

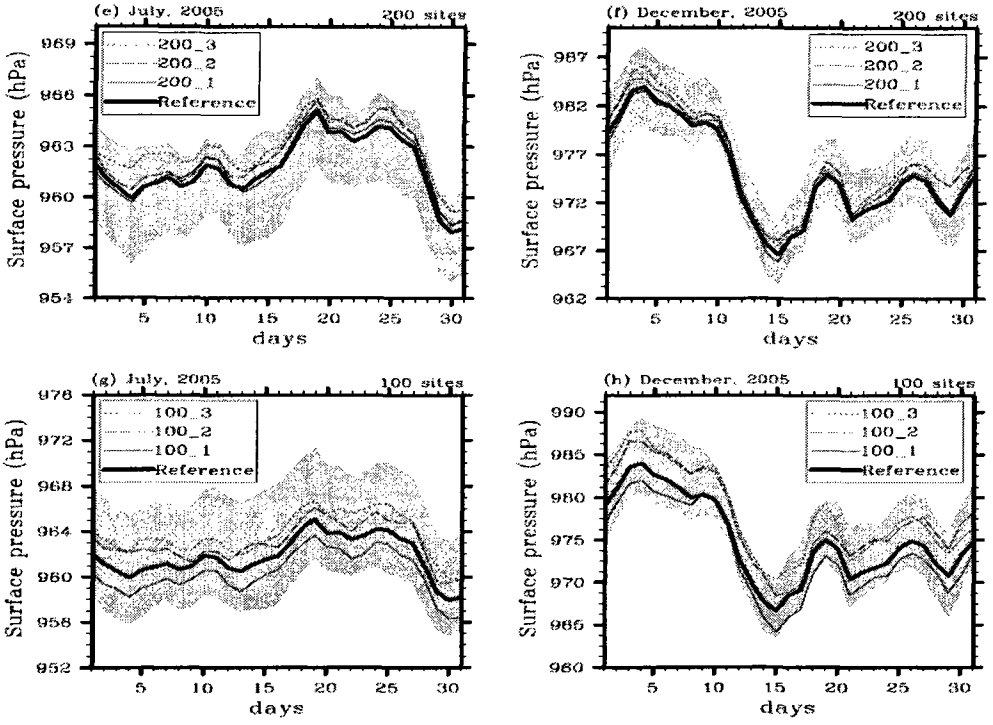
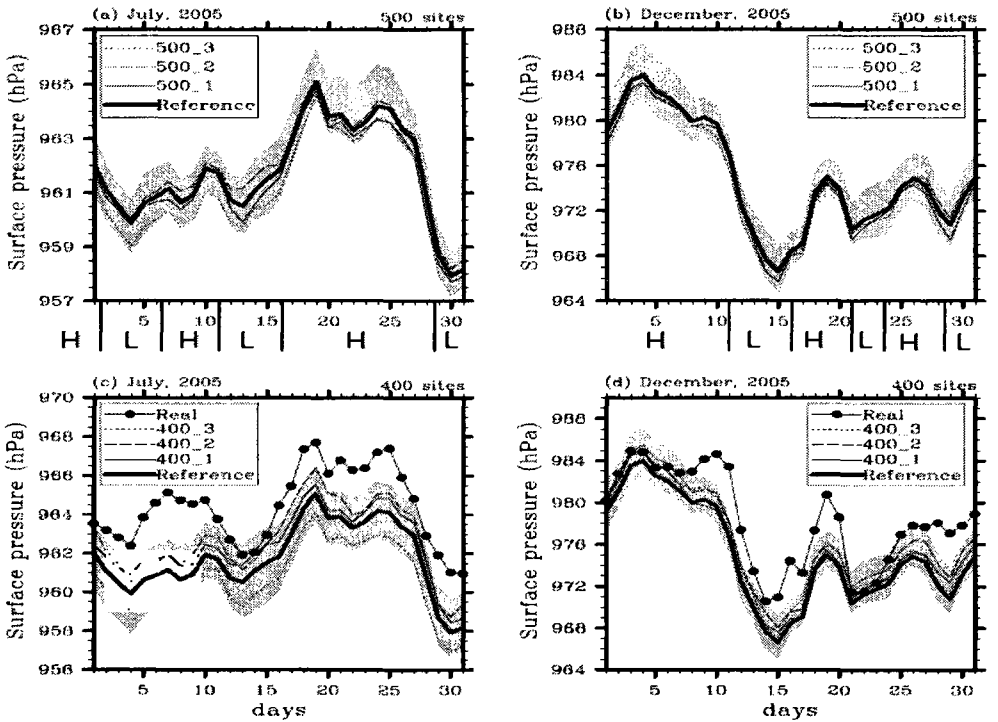


Fig. 5.4 Temporal behavior of regional averages of sea-level pressure as obtained for the reference data using all data within all $2.8^{\circ} \times 2.8^{\circ}$ areas and as derived for various 500-


(a) July, 2005
500 sites
500_3
500_2
500_1
Reference
Surface pressure (hPa)
H
L
(b) December, 2005
500 sites
(c) July, 2005
400 sites
Real
400_3
400_2
400_1
Reference
days
(d) December, 2005
400 sites

sites-networks based on the "sites" within the $2.8° \times 2.8°$ areas for (a) July, (b) December, for various 400-sites-networks for (c) July, and (d) December, for various 200-sites-networks for (e) July and (f) December, (e) like for (a), but for July, (f) December, and for various 100-sites-networks in (g) July and (h) December. In (c) and (d) the solid line with filled circles represents the regional averages derived from the real network (411 sites). In all panels, all other lines represent the regional averages with lowest error values among the ten setups of the respective network of given density; the shaded regions represent the maximum over- and underestimation of the reference regional averages found for the ten networks of same density. The letters H and L in (a) and (b) represent the days with high and low pressure situation. Data shown are for 2005, but 2006 and 2007 show similar general behavior with respect to differences among networks of same density, the range of over- and underestimation by networks and among networks of different density as well as reaction to events (e.g. frontal passages). See text for further details.

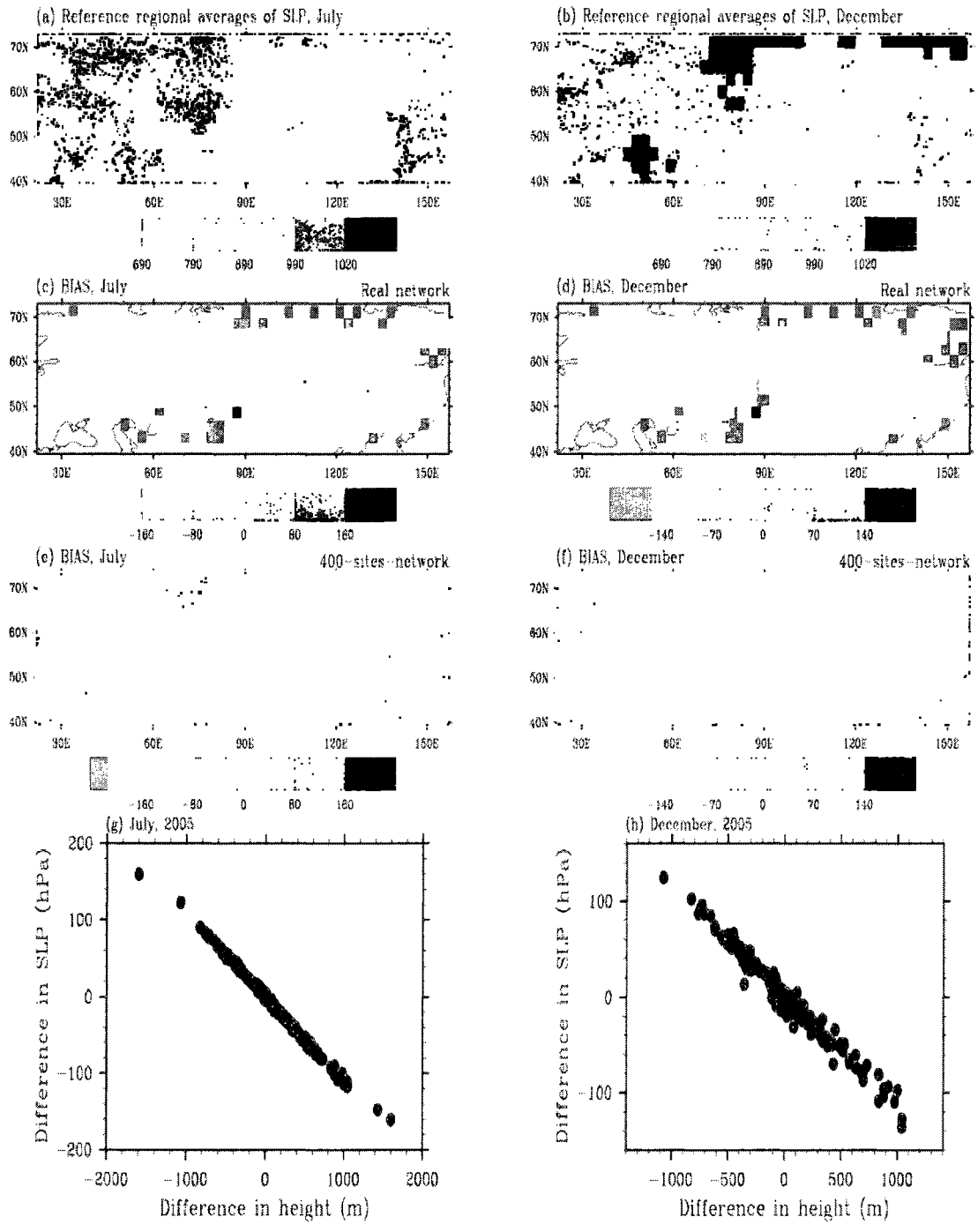


Fig. 5.5 Spatial distribution of regional averages of sea-level pressure in hPa for (a) July, (b) December, biases for (c) July, (d) December between regional averages of sea-level

pressure estimated from the real network and the reference network, biases for (e) July, (f) December between regional averages of sea-level pressure estimated from the ten 400-sites networks and the reference network and scatter plot of spatial differences between regional averages of sea-level pressure estimated from the real network and the reference network and spatial differences between terrain elevation (m) estimated from the real network and the reference network for (g) July and (h) December. In parts (e) and (f), always the maximum values for the 10 400-sites networks is shown when several networks have values for the same $2.8^{\circ} \times 2.8^{\circ}$ area. Since geographical trends are marginal for all artificial networks as demonstrated by parts (e) and (f) spatial plots for artificial networks are not presented any further. Data shown are for 2005. The general distribution of errors looks similar for 2006 and 2007 (therefore not shown). Legends differ among panels. White areas in the plot represent areas with no site in the real network. Note that no interpolated values are shown to avoid the mix of differences truly due to network density and design (shown here) and those from interpolation methods that are not in the focus of this paper (therefore not discussed).

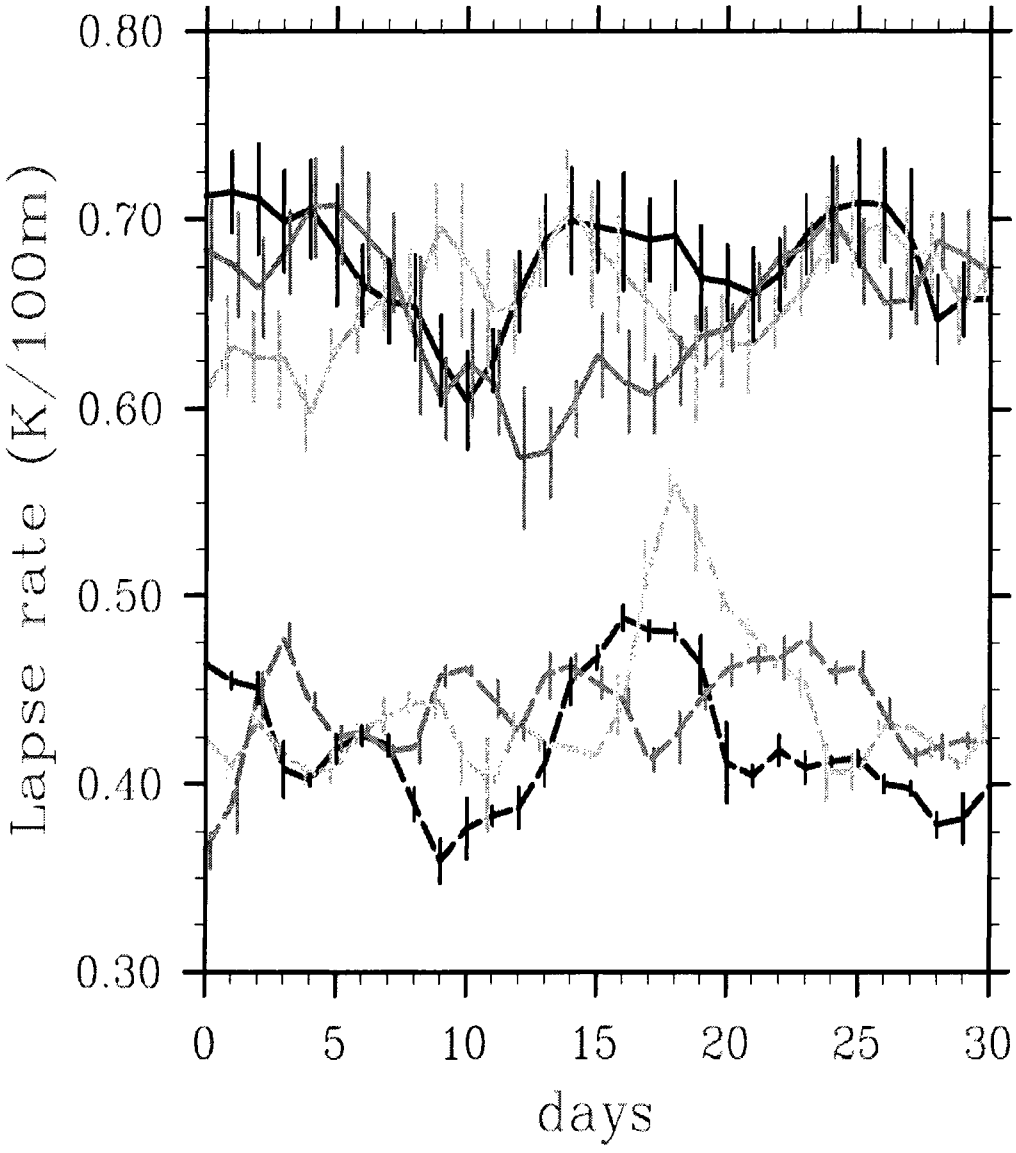


Fig. 5.6 Temporal evolution of the reference regional average environmental lapse rate for 2005 (black line), 2006 (dark gray line) and 2007 (light gray line) for July (solid lines) and December (dashed lines). The bars indicate the temporal and spatial average standard deviation of the environmental lapse rate on a given day.

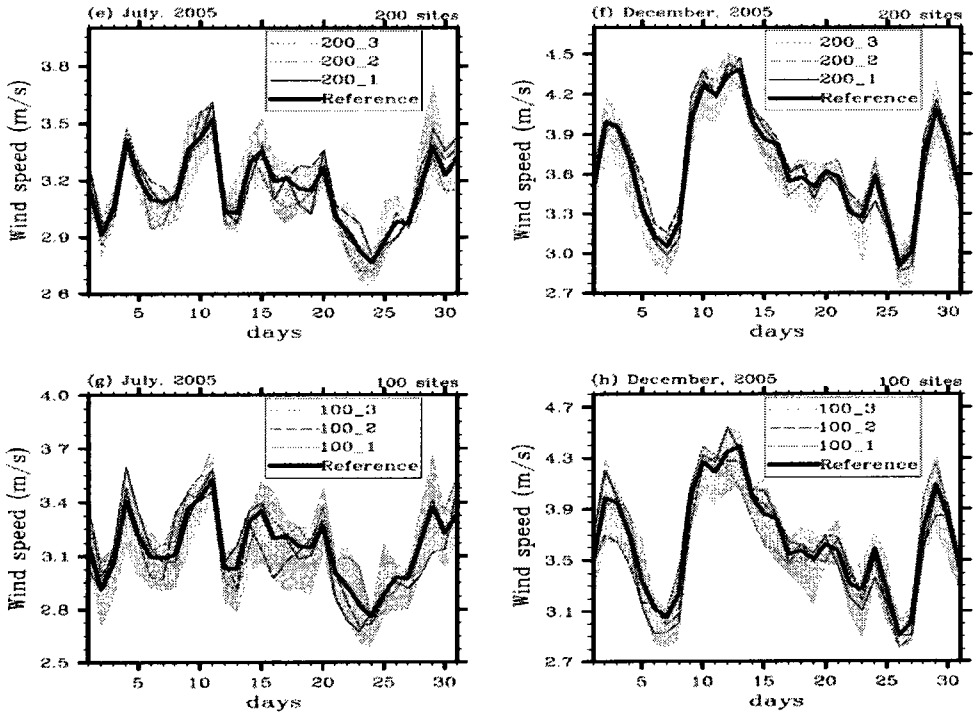
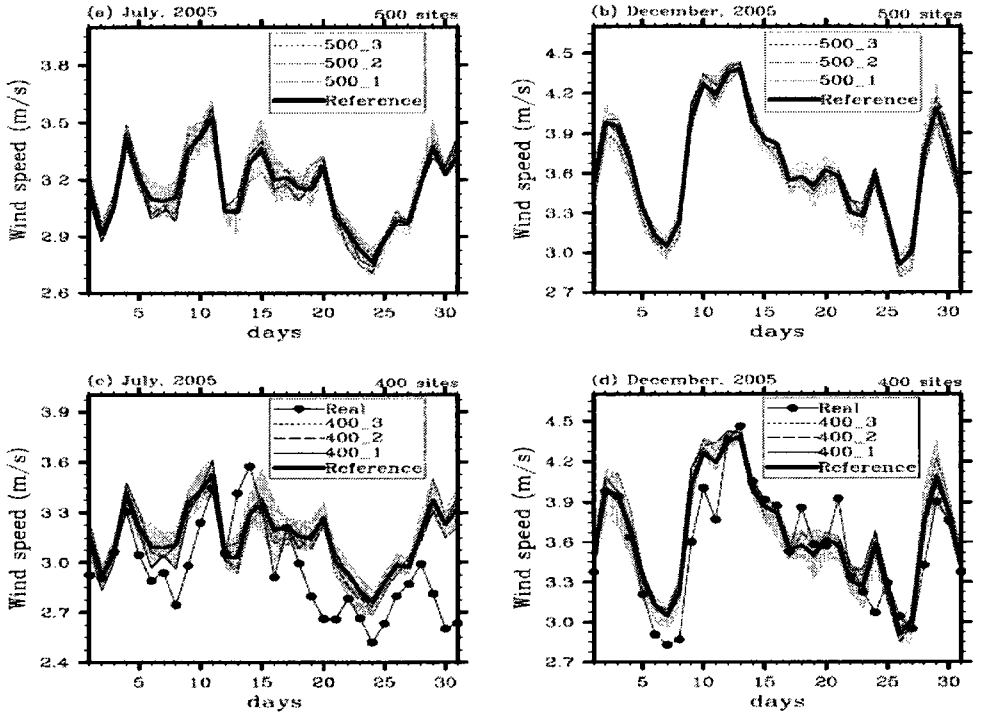


Fig. 5.7 Like Fig. 5.4, but for 10m-wind-speed in m/s


(a) July, 2005
500 sites
500_3
500_2
500_1
Reference
Wind speed (m/s)
days
(b) December, 2005
500 sites
500_3
500_2
500_1
Reference
Wind speed (m/s)
days
(c) July, 2005
400 sites
Real
400_3
400_2
400_1
Reference
Wind speed (m/s)
days
(d) December, 2005
400 sites
Real
400_3
400_2
400_1
Reference
Wind speed (m/s)
days

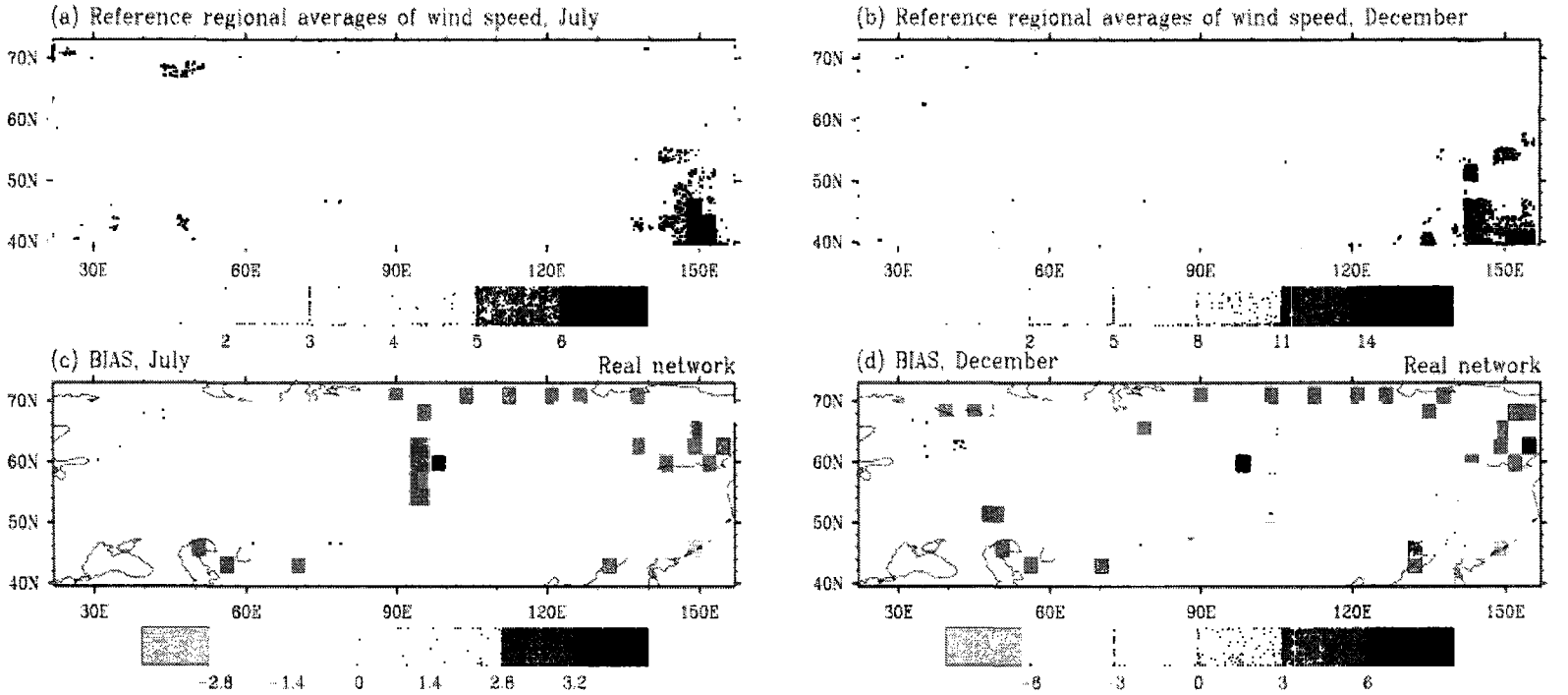


Fig. 5.8 Spatial distribution of regional averages of 10m-wind-speed in m/s for (a) July, (b) December, biases for (c) July, (d) December between regional averages of 10m-wind-speed estimated from the real network and the reference network.

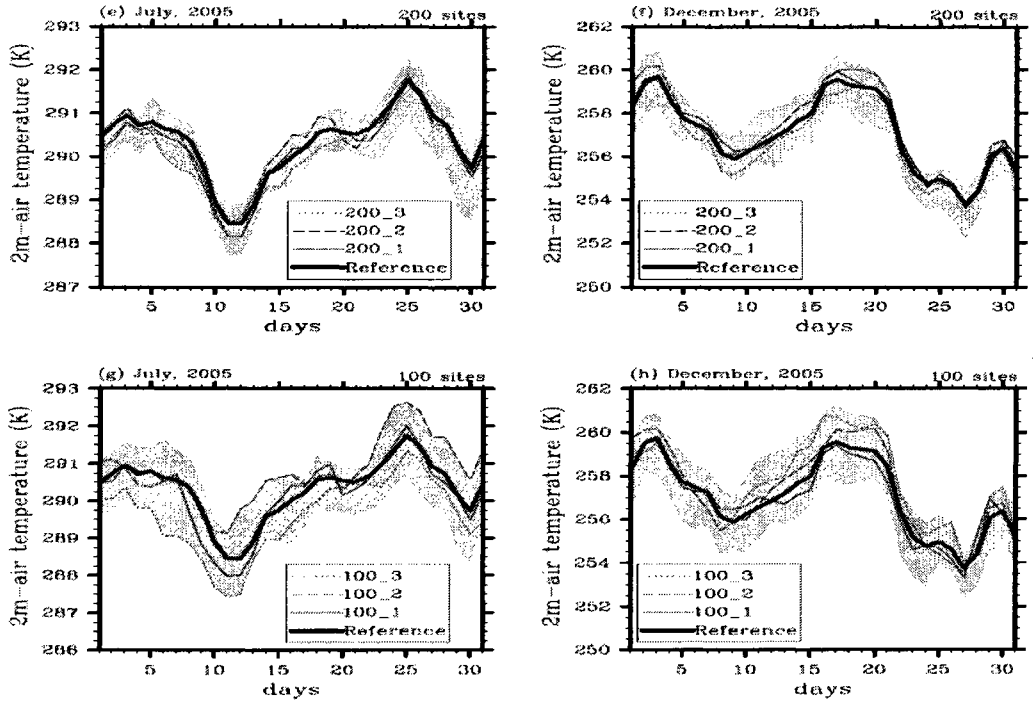
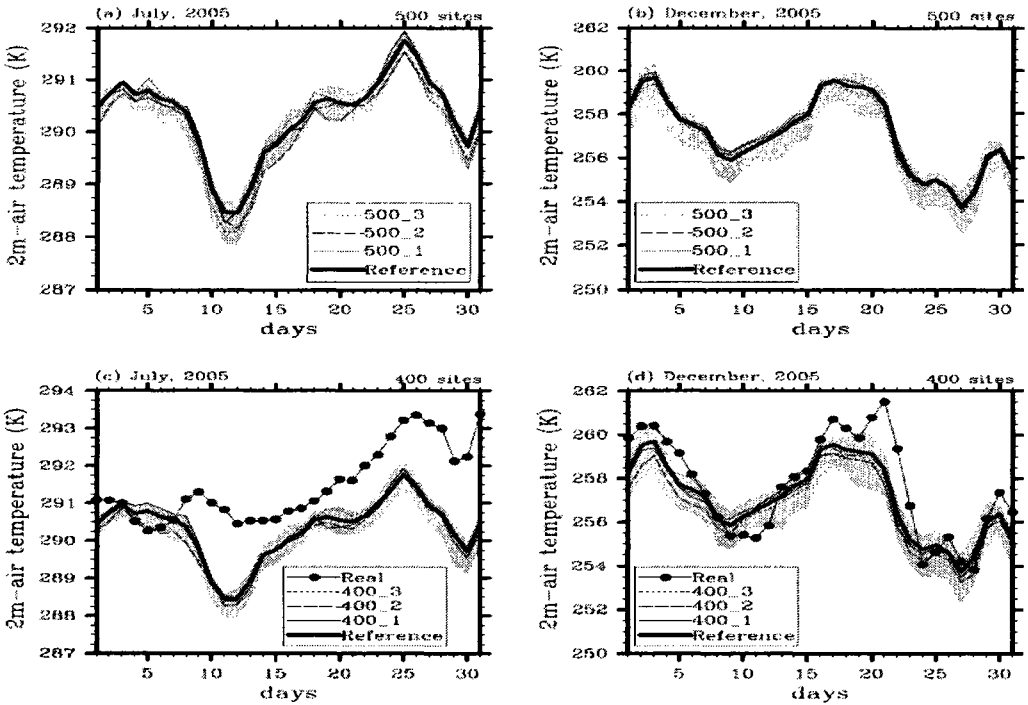


Fig. 5.9 Like Fig. 5.4, but for 2-m temperature in K.


(a) July, 2005
500 sites
2m-air temperature (K)
days
500_3
500_2
500_1
Reference
(b) December, 2005
500 sites
(c) July, 2005
400 sites
Real
400_3
400_2
400_1
Reference
(d) December, 2005
400 sites

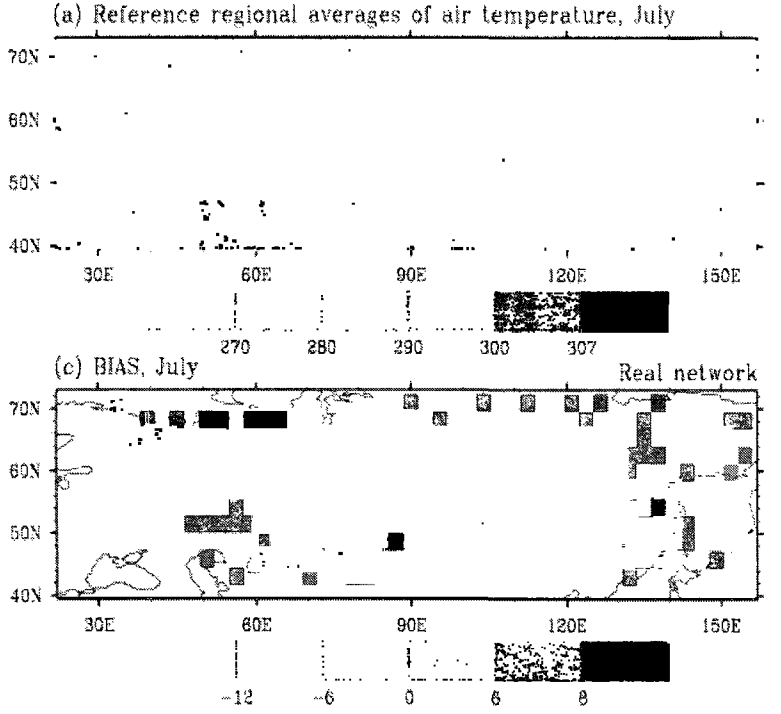
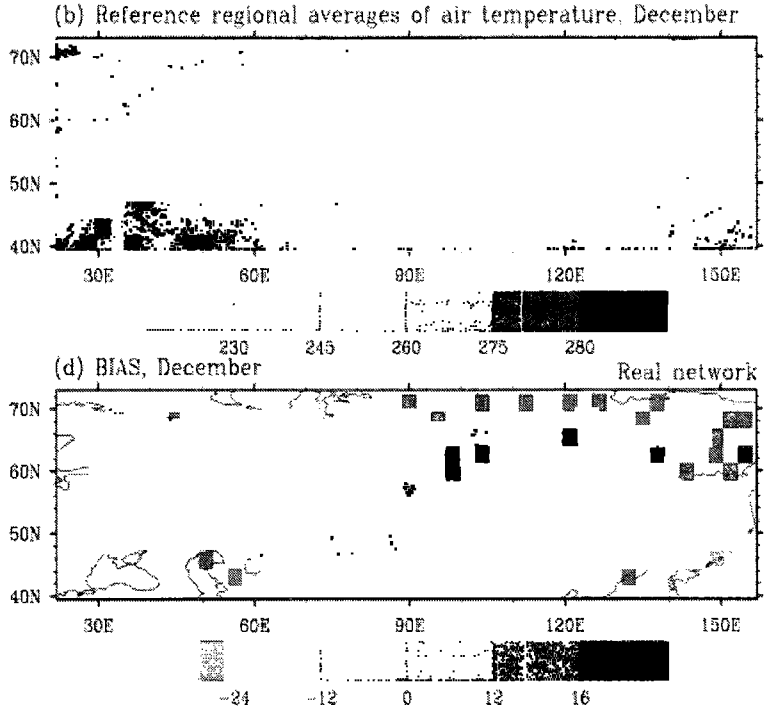


Fig. 5.10 Like Fig. 5.8, but for 2-m temperature in K.


(b) Reference regional averages of air temperature, December
70N
60N
50N
40N
30E
60E
90E
120E
150E
230
245
260
275
280
(d) BIAS, December
Real network
70N
60N
50N
40N
30E
60E
90E
120E
150E
-24
-12
0
12
16

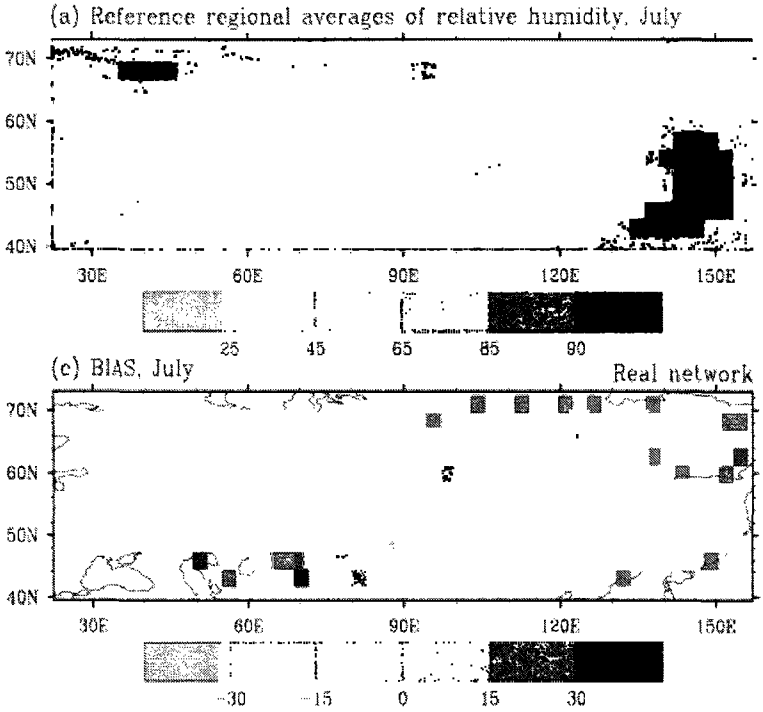
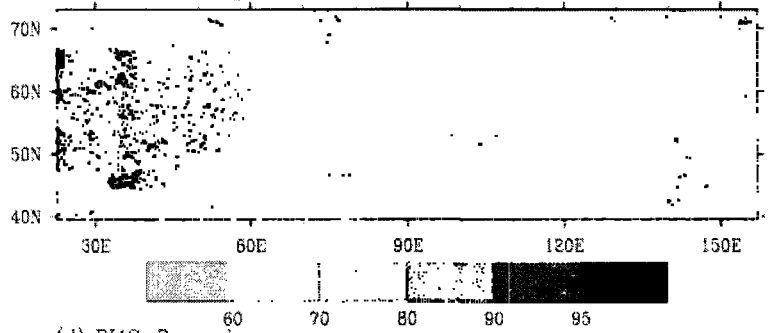
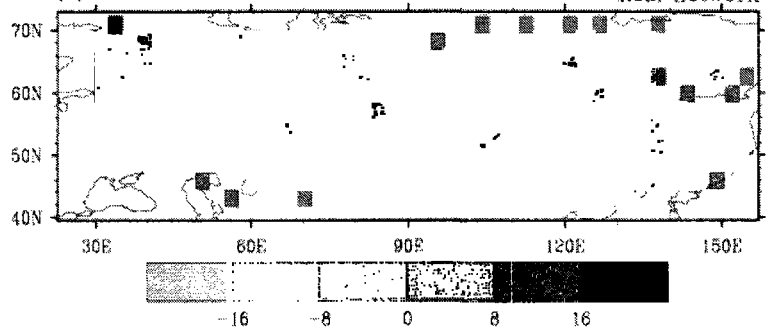


Fig. 5.11 Like Fig. 5.8, but for relative humidity in %.

(b) Reference regional averages of relative humidity, December



(d) BIAS, December Real network



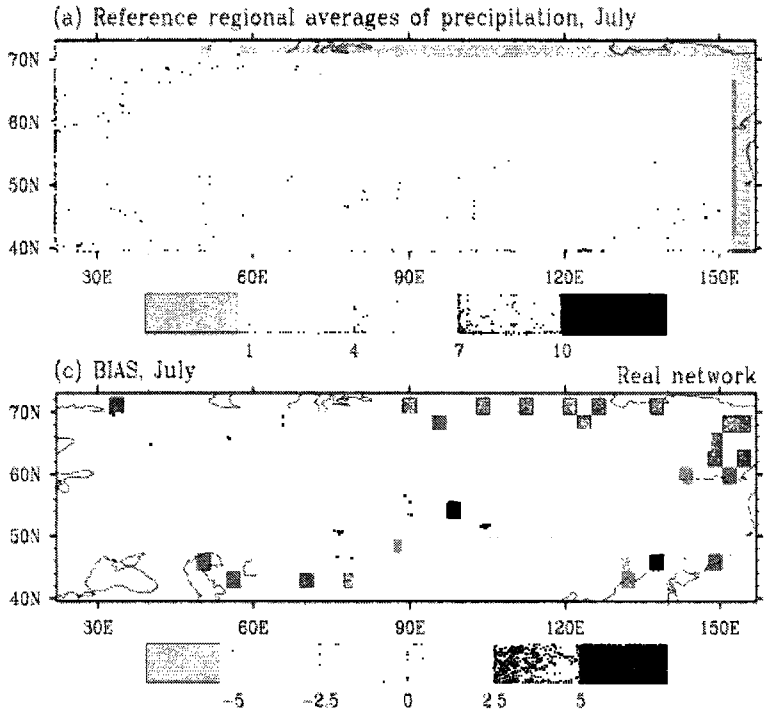
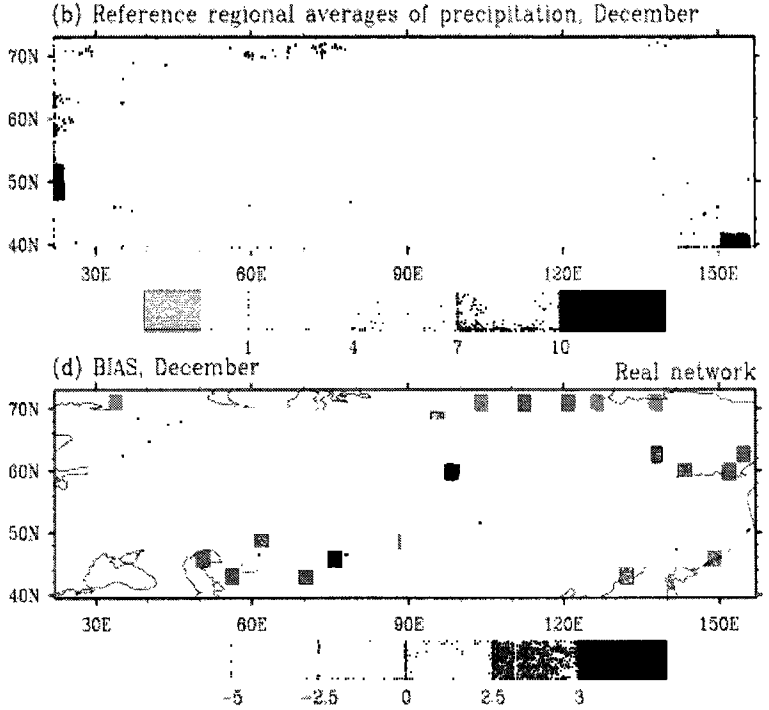


Fig. 5.12 Like Fig. 5.8, but for precipitation in mm/d.


(b) Reference regional averages of precipitation, December
70N
60N
50N
40N
30E
60E
90E
120E
150E
1
4
7
10
(d) BIAS, December
Real network
70N
60N
50N
40N
30E
60E
90E
120E
150E
-5
-2.5
0
2.5
3

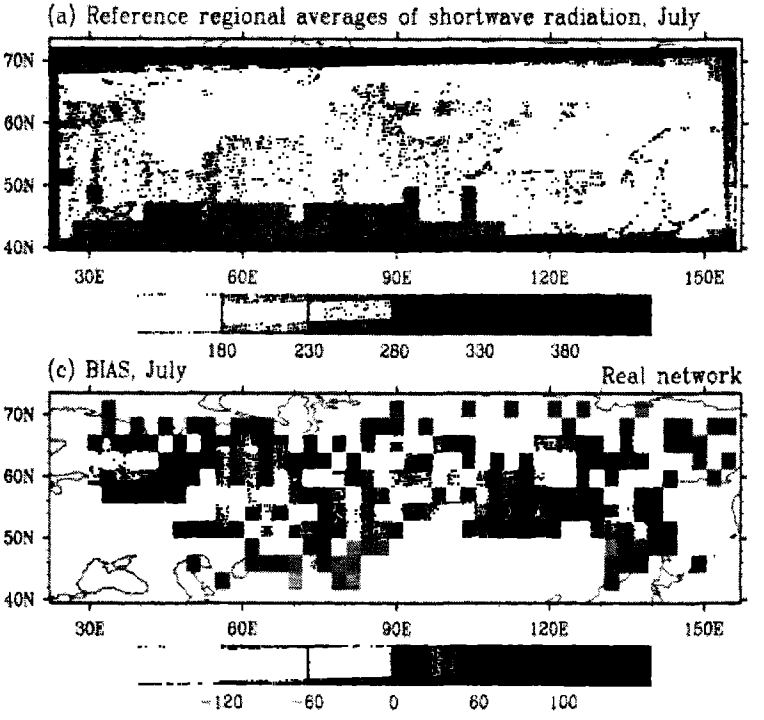
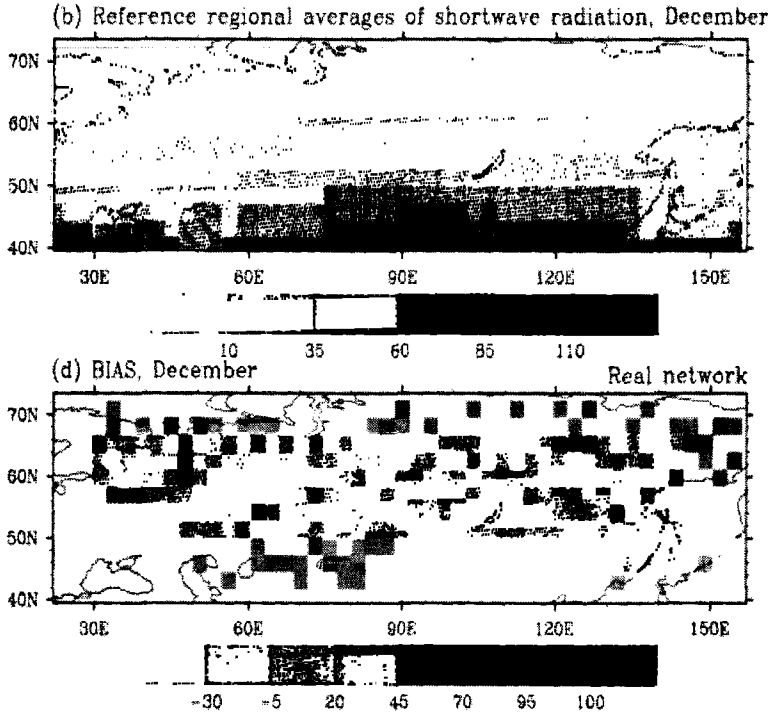


Fig. 5.13 Like Fig. 5.8, but for shortwave radiation in Wm^{-2}.


(b) Reference regional averages of shortwave radiation, December
70N
60N
50N
40N
30E
60E
90E
120E
150E
10
35
60
85
110
(d) BIAS, December
Real network
70N
60N
50N
40N
30E
60E
90E
120E
150E
-30
-5
20
45
70
95
100

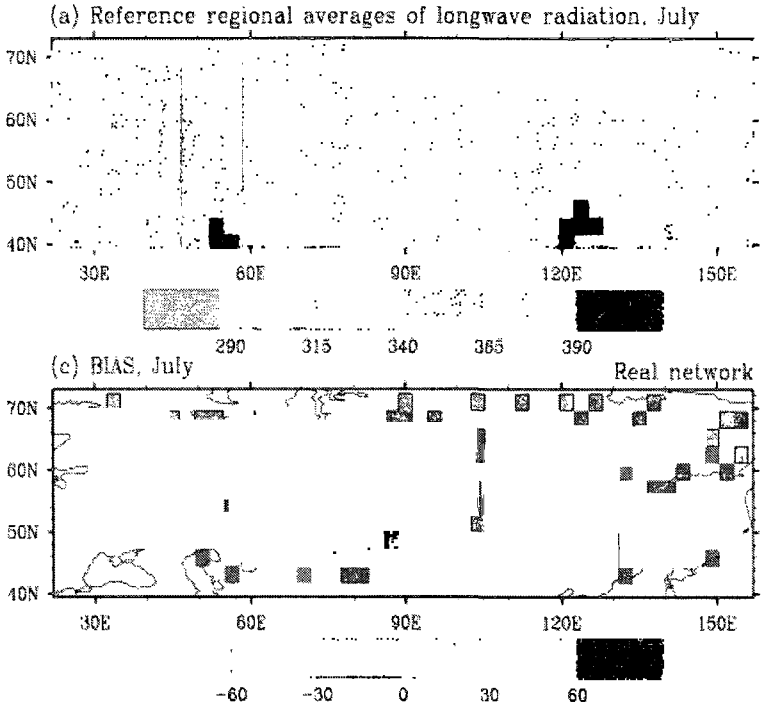
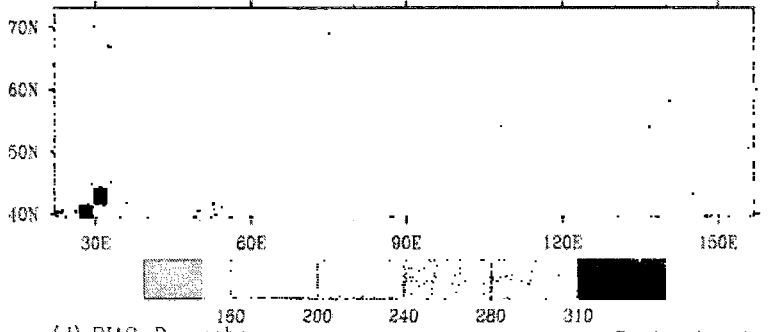
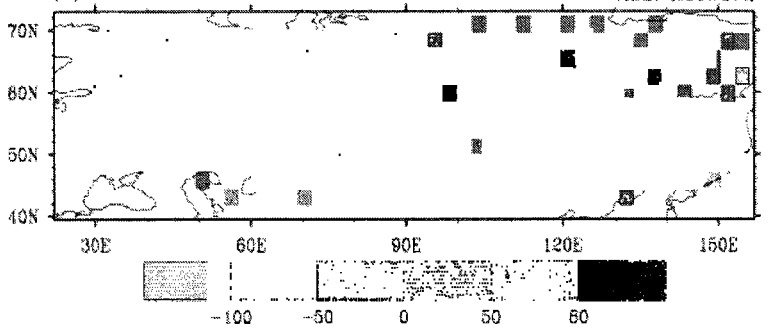


Fig. 5.14 Like Fig. 5.8, but for long-wave radiation in Wm^{-2}.

(b) Reference regional averages of longwave radiation, December



(d) BIAS, December Real network



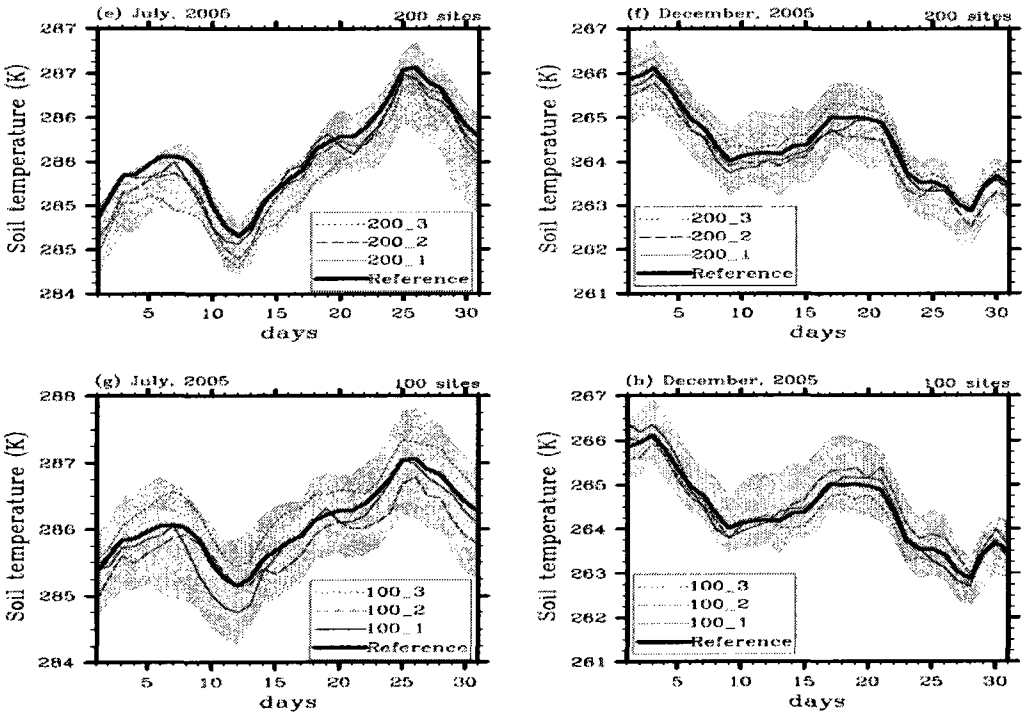
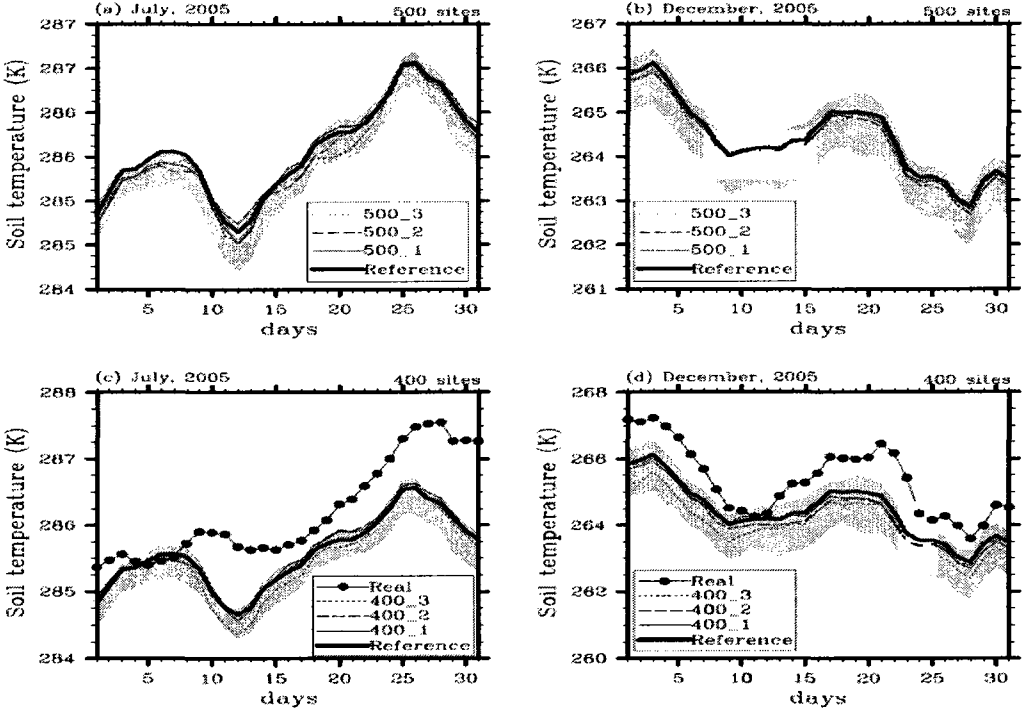


Fig. 5.15 Like Fig. 5.4, but for soil-temperature at 0.2 m depth in K.


(a) July, 2005
500 sites
Soil temperature (K)
days
500_3
500_2
500_1
Reference
(b) December, 2005
500 sites
Soil temperature (K)
days
500_3
500_2
500_1
Reference
(c) July, 2005
400 sites
Soil temperature (K)
days
Real
400_3
400_2
400_1
Reference
(d) December, 2005
400 sites
Soil temperature (K)
days
Real
400_3
400_2
400_1
Reference

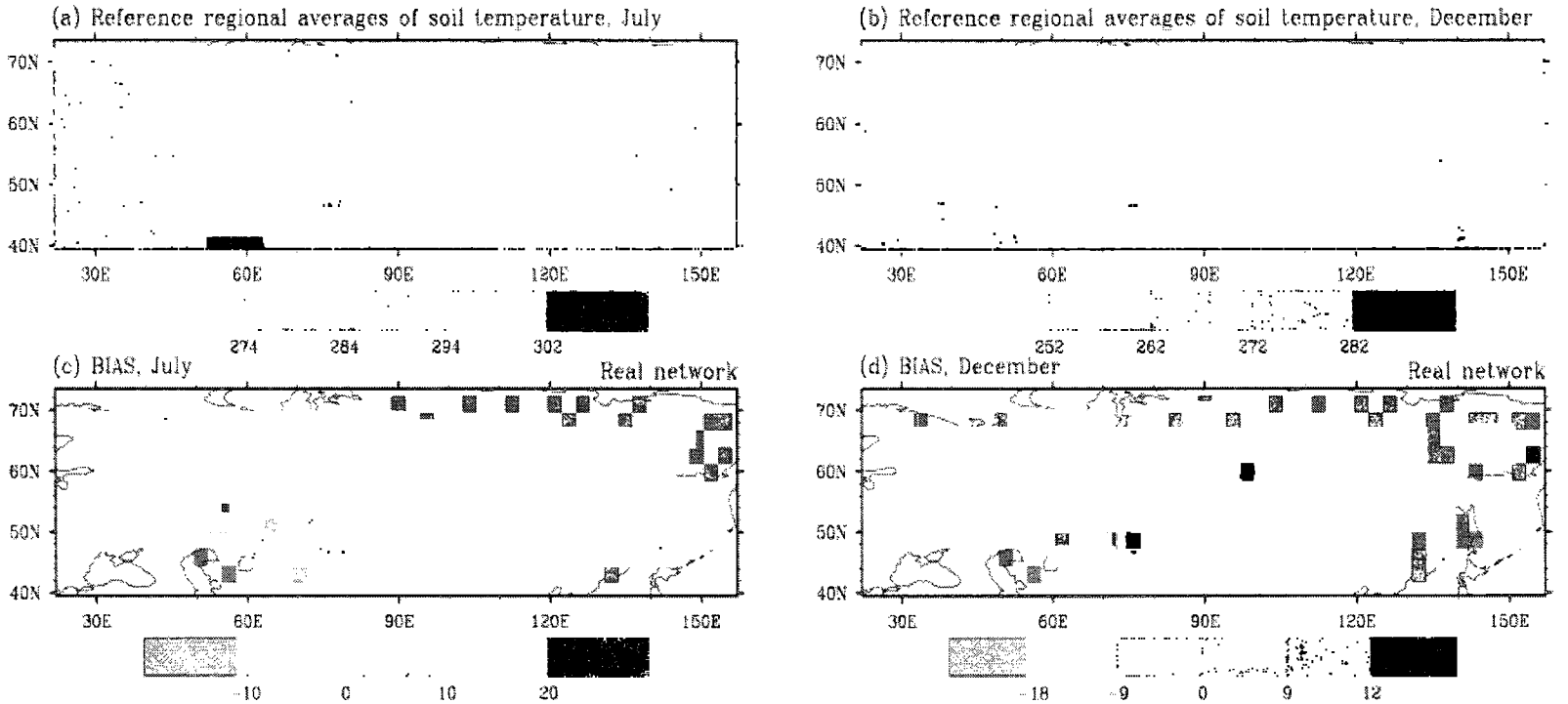


Fig. 5.16 Like Fig. 5.8, but for soil-temperature at 0.2 m depth in K.

Chapter 6 Sources of discrepancy between CCSM simulated and gridded observation-based soil-temperature over Siberia: The influence of site density and distribution**

Abstract

Soil-temperature climatologies determined at different depths from simulations with the Community Climate System Model version 3 (CCSM) capture the annual phase of gridded soil-temperature climatologies based on observations for 1951-1980, 1961-1990, and 1971-2000, but not the amplitude; some of these discrepancies can be attributed to simulated forcing (PaiMazumder et al. 2008). By using soil-temperature data simulated by Weather Research and Forecasting (WRF) model it is shown that some of the discrepancies between CCSM-derived and gridded observed climatologies may result from the interpolation required for gridding and/or network design (density and distribution of sites).

6.1 Introduction

Accurate simulation of soil temperature in Climate System and Earth System Models is essential because soil temperature influences high-latitude hydrology, biochemical processes and ecosystems. Soil temperatures are mostly controlled by the surface water and energy balance, which explains the strong connection and feedback between soil and near-surface atmospheric conditions. In the Arctic and Subarctic, the onset, duration, thickness, density and structure of seasonal snow-cover strongly influence soil temperatures (e.g. Zhang et al. 1996, Mölders and Romanovsky 2006).

Soil temperatures simulated at different depths by the Community Climate System Model version 3 (CCSM; Collins et al. 2006a) are evaluated over Siberia for three climatologies (1951-1980, 1961-1990, 1971-2000) by means of observational data

---

** PaiMazumder D, Mölders N (2008) Sources of discrepancy between CCSM simulated and gridded observation-based soil-temperature over Siberia: The influence of site density and distribution. 9th International Conference on Permafrost (NICOP):1351- 1356

(PaiMazumder et al. 2008) provided by National Snow Ice Data Center (NSIDC; Zhang et al. 2001). PaiMazumder et al. (2008) also evaluated CCSM derived climatologies of near-surface temperature, cloud fraction, precipitation and snow depth with those from ERA40 reanalysis, International Satellite Cloud Climatology project (ISCCP), Global Precipitation Climatology Center (GPCC) and NSIDC, respectively, to examine the sources for discrepancies between simulated and observed soil-temperature climatology. Inaccurate simulation of near-surface temperature, cloud fraction, precipitation and snow depth may have some influences on discrepancies between CCSM-derived and observed soil-temperature climatology, but do not explain all discrepancy found; sensitivity studies with slightly altered plant functional types and percentage of sand attributed marginal discrepancies from incorrect percentages of sand and/or plant types (PaiMazumder et al. 2008).

Typically climate models like CCSM provide soil temperatures that represent a volume average of several 100 square kilometers in horizontal extension of several centimeters in thickness. It is obvious that soil temperatures simulated for such a volume are difficult to compare to measurements at a site (point measurements). Therefore, it has become common practice to interpolate available measurements to the grid of the climate model (e.g. Li et al. 2008). It is obvious that such interpolation may introduce uncertainty into the grid-cell averages and, hence, the evaluation. Since the gridded soil-temperature climatologies are based on measurements projected onto the CCSM3-grid by Cressman-interpolation (PaiMazumder et al. 2008), some discrepancies between CCSM-derived and gridded observed climatologies may result from interpolation and/or network density and distribution. Observational networks are often designed with accessibility and ease of maintenance in mind. Most of the Siberian soil-temperature sites are long term agricultural monitoring stations. Consequently, the observational network follows agricultural-used land along major haul-ways and is not uniformly distributed. Hence, the density and/or design of network may bias the regional averages estimated there from. Mitchell et al. (2004) assessed accuracy and reliability of gridded data and concluded that observed gridded data (1901-2000) are not appropriate for climate change. They also

describe the development of high resolution (0.5°) of gridded data set (Climate Research Unit (e.g. CRU TS 2.0) data) for the globe derived from climatological observations and transient coupled atmosphere-ocean general circulation model (GCM) simulations; the gridded dataset depends on the applied interpolation algorithms and always has to be associated with an assessment of the accuracy of the grid point values. Therefore, it is essential to assess the uncertainty in regional averages resulting from the density and/or design of an observational network. The aim of our case study is to exemplarily investigate this uncertainty to further assess the discrepancies between CCSM-derived and observed soil-temperature climatologies found by PaiMazumder et al. (2008). In doing so, the Weather Research and Forecasting (WRF; Skamarock et al. 2005) model is used to provide a dataset of soil temperatures that will be considered as "reference" for determination of regional averages to which data from a real network and artificial networks are compared to assess the accuracy of gridded datasets based on station data and to develop recommendations for network design to optimize their use for model validation.

6.2 Experimental design

6.2.1 Brief model description

The CCSM is a fully coupled climate model to simulate the Earth system over broad ranges of spatial and temporal resolutions. It consists of the Climate Atmospheric Model version 3 (CAM3; Collins et al. 2006b), the Community Land Model version 3 (CLM3; (Dai et al. 2003, Oleson et al. 2006), the Community Sea Ice Model version 5 (CSIM5; Briegleb et al. 2004) and the parallel Ocean Program version 1.4.3 (POP; Smith et al. 1992). These four components exchange data via a coupler without flux correction.

CCSM is run with 26 vertical layers at a spectral truncation of T42 corresponding to a spatial resolution of ≈2.8°x2.8°. CCSM is started with the ecliptic conditions of 1-1-1950 and CO_2 concentration of 355 ppmv. Each model component is spun up separately. Based on these simulations, we determine three climatologies, 1951-1980, 1961-1990, and 1971-2000.

The WRF is a mesoscale non-hydrostatic model. Out of the variety of physical options we use the following model setup: Cloud formation and precipitation processes at the resolvable scale are considered by Thompson et al.'s (2004) five-water class (cloud-water, rainwater, ice, snow, graupel) mixed-phase bulk-microphysics parameterization. The Grell-Devenyi (2002) ensemble parameterization considers subgrid-scale convective clouds. The Goddard shortwave-radiation scheme and the Rapid Radiative Transfer Model (Mlawer et al. 1997) are applied. The Yonsei University scheme (Skamarock et al. 2005) is used for simulating atmospheric boundary layer processes. Monin-Obukhov similarity theory is applied for surface-layer physics. Soil temperature, volumetric ice and water content, snow temperature and density and the exchange of heat and moisture at the land-atmosphere interface are determined by a modified version of the Rapid Update Cycle land-surface model (Smirnova et al. 1997, 2000).

The WRF domain encompasses Siberia by 70x150 grid-points with a grid-increment of 50 km and 31 vertical layers from the surface to 50 hPa. Soil conditions are determined at six levels. In the presence of snow, five snow layers are considered. The time step is 200 s. The National Centers for Environmental Prediction (NCEP) 1.0^{o}x1.0^{o} and 6h-resolution global final analyses (FNL) serve as initial and boundary conditions. For our case study, we perform simulations for July and December 2005. They start daily at 1800 UT for 30 hours of integration. We discard the first six hours as spin-up time.

6.2.2 Analysis

To estimate uncertainty due to network density and design, WRF-simulated soil temperatures serve to represent data from an optimal, dense, and equally distributed observational network. Regional averages of soil temperatures determined from the WRF output for July and December, 2005 are considered to be the "reference".

Regional averages of soil temperature are determined for the 411 sites of the actual historic observation network used in PaiMazumder et al.'s (2008) CCSM soil-temperature evaluation. Herein, the soil temperature simulated for a WRF-grid cell wherein a site falls is taken as the soil temperature for that site. This procedure is

common practice in mesoscale modeling (e.g. Narapusetty and Mölders 2005). Four artificial networks are assumed with 500, 400, 200, and 100 arbitrarily taken WRF grid-cells as "sites". These networks are denoted 500-, 400-, 200- and 100-site-networks hereafter. Soil temperatures obtained from WRF simulations at the 500, 400, 200 and 100 sites are used to calculate the regional averages for these networks. These regional averages are compared with the reference regional averages to assess the contribution of network density and design to uncertainty in gridded data used for evaluation of climate model data.

Since systematic and nonsystematic errors can contribute to any simulation result as well as to regional averages obtained from different networks, performance measures like bias, standard deviation of errors (SDE), root mean square errors (RMSE) and correlation coefficients (e.g. Anthes 1983, Anthes et al. 1989) are calculated at different spatial and temporal scales for the various networks. The performance measures and correlation coefficients are determined to evaluate the discrepancies between the regional averages obtained from the "reference" and those of a network. They are calculated for all networks for the daily and monthly course.

To estimate the uncertainties in regional averages resulting from the density and design of networks, we compare the regional averages of soil temperature obtained from the WRF simulation ("reference") and various site networks (500-, 400-, 200-, 100-site-network and historic networks). In this case study, we consider the accuracy of soil temperature measurements to be within ±0.5 K for the reasons discussed in PaiMazumder et al. (2008). The regional averages of soil temperature estimated from the "reference" and the different networks will be considered to be in good agreement if regional averages obtained from the different networks lie within the above mentioned uncertainty range to the "reference".

In a next step, we compare the uncertainty determined as described above with the discrepancies found between CCSM-derived and gridded observation derived climatologies by PaiMazumder et al. (2008) to assess how much the network design may explain some of these discrepancies.

6.3 Results

6.3.1 Impact of network design on regional averages

Regional averages of soil temperature obtained from the artificial networks are highly correlated with those of the reference (>0.972) at all depths in both months. Regional averages of soil temperature obtained from the historic network are higher correlated with the "reference" in December (R=0.921) than in July (R=0.732) at 00 UT, 06 UT, 12 UT, 18 UT at all depths.

In both months, the daily spatial standard deviations of soil temperatures obtained from the historic network are higher than those of all other networks. The standard deviations of soil temperature obtained from the various networks are higher at all depths in December than in July. In December, in general, soil temperatures vary strongly in space due to the large horizontal differences in snow cover and/or thickness. Thus, taking measurement along the haul ways leads to larger standard deviations of the regional averages in winter than summer because of snow conditions and terrain height of the site. In July, for given insolation and soil type, soil heating/cooling varies less in space than in December when differences in snow-cover/thickness may strongly affect soil temperatures. Consequently, taking the measurement along a haul way has less impact on the regional average and its standard deviation for Siberia in summer than winter.

At 0.2m depth, soil temperatures obtained from the historic network overestimate the reference average by up to 1.5 K and 1.8 K in July and December, respectively (Fig. 6.1). The historic network also fails to capture the timing of the soil-temperature maxima and minima represented in the reference average. For example, on the 11^{th} and 26^{th} July, upper soil temperatures from the historic network average do not reflect the warm periods seen in the reference average (Fig. 6.1). These differences in timing of extremes between the regional averages of the historic network and "reference" occur due a frontal system passing Siberia. The non-equal distribution of sites of the historic network therefore gets "biased" to the time when the system passes the majority of the sites. In December, the high bias found for the historic network may partly be explained by the fact that the sites of the historic network may not well represent the regional differences in snow cover

and/or thickness. The 100-site-network also fails to capture the regional soil-temperature averages obtained from the "reference" with 0.8 K bias in July (Fig. 6.1), while in December, 100-site-network captures the reference average well. Obviously, a randomly distributed 100-site-network represents soil conditions in winter well when the soil is partly insolated by snow. However, in summer, convection may lead to spatial differences in soil heating due to shading by clouds and/or heat-input by precipitation that a 100-site-network cannot capture appropriately. Regional averages of soil temperature obtained from 500-, 400- and 200-site-network provide acceptable results in comparison to the reference average at all depths in both months. The historic network shows higher biases and RMSEs than all other networks in both months (Fig. 6.1, Table 6.1). This means that the historic network introduces some bias into regional averages and any evaluation study therewith.

Higher systematic bias may have occurred due to the difference in landscape and terrain elevations between the regional averages derived from the "reference" and historic network. At 0.2 m depth (0.4 m and 1.6 m) RMSEs for the historic network reach up to 1.5 K (1 K and 1.9 K) and 1.8 K (1.3 K and 1.1 K) in July and December, respectively (Fig. 6.1). Errors in regional soil-temperature averages based on the historic network are also high for the upper soil layer and errors decrease at 0.4 m depth and increase again for deeper soil layers in both months. The high diurnal variability close to the surface may cause the high errors in the upper soil (Table 6.1). The increase found for deeper soil layers may be related to the constant lower boundary condition used in WRF. These soil-temperature values are from climatologies which differ notably in space. Thus, the historic network cannot represent the "reference" because the sites are not randomly distributed to capture the regional pattern.

6.3.2 Evaluation of CCSM by gridded data

As reported by PaiMazumder et al. (2008), CCSM captures the phase of the 30-years average annual soil temperature curves well at all depths, but not the amplitude. CCSM overestimates the 0.2-m soil temperature for the majority of the grid cells over Siberia

from December to March for the first climatology (1951-1980) (Fig. 6.2). Similar is true for 0.4 m and 1.6 m depths, but with marginally decreasing frequency with increasing depth, whereas at 0.8 m and 3.2 m, CCSM overestimates soil temperature throughout the year (Fig. 6.2). In April, soil temperature will be overestimated for most of the grid-cells at all depths if soil temperature is below freezing and underestimated otherwise, leading to overall overestimation. In May and June, the general pattern shifts towards underestimation at all depths except 3.2 m (Fig. 6.2). In July and August, CCSM tends to underestimate soil temperature by up to 1.2 K at 0.2 m and overestimates them by up to 0.5 K at 0.4 and 0.8 m depth. At 1.6 m, the tendency to overestimate soil temperature is obvious for the colder (<275 K), but less obvious for the warmer (>285 K) end of the temperature range. In September, soil temperatures are underestimated at 0.2 and 0.4 m depth by up to 0.6 K and 0.4 K for most grid-cells (Fig. 6.2). At 0.8 m and 1.6 m, CCSM overestimates soil temperatures by up to 0.4 K and 1.1 K, respectively, for most grid-cells. In October, the general pattern again shifts towards overestimation by up to 2.5 K (Fig. 6.2). In November, simulated soil conditions are too warm for soil temperatures below the freezing point at all depths. PaiMazumder et al. (2008) also found that biases and RMSEs decrease with increasing depth because most variability occurs near the surface. RMSEs are higher in winter than in the other seasons for all three climatologies. Mean annual soil temperatures are over-estimated by 2.5 K, on average.

Overall, CCSM simulates the annual average soil temperature reasonably well though its performance is better in summer than in winter. Due to acceptable RMSEs and high correlations (cf. Fig. 6.2), PaiMazumder et al. (2008) concluded that the fully coupled CCSM acceptably simulates soil temperature. As pointed out by these authors, CCSM also has some difficulties to capture near-surface temperature, cloud fraction, precipitation and snow depth with biases (RMSEs) -1.0 K (3 K), 0.32 % (0.52 %), 7.6 mm/month (19.9 mm/month) and 0.04 m (0.09 m), respectively (for further details see Table 1 in PaiMazumder et al. 2008). Hence the inaccurate simulation of near-surface temperature, cloud fraction, precipitation and snow depth may contribute to the discrepancies between CCSM-derived and observed soil temperature climatology.

Difference in plant functional types between model and real world and decreasing and increasing sand percentage in the model may marginally affect soil temperature. Thus, incorrect assumptions on the mineral soil type cannot explain the discrepancies found. Another source of these discrepancies may be that CCSM only considers mineral soils. However, large areas of Siberia have organic soils at least in the upper soil layers. The thermal and hydraulic properties of mineral and organic soils differ strongly and yield to appreciable differences in soil temperature and soil-water freezing behavior (cf. Mölders and Walsh 2004, Lawrence and Salter 2007). Nicolsky et al. (2007) showed that incorporating organic matter in CLM3 significantly changes the soil temperature simulation. Thus, providing gridded data of organic material distribution and consideration of organic material in CCSM are essential future steps for the scientific community to take.

6.4 Discussion and conclusions

PaiMazumder et al. (2008) found that CCSM tends to overestimate soil temperature in winter and underestimate in summer with better performance in summer than winter. In spring and fall, simulated and observed climatologies agree the best. Therefore, we performed a case study with WRF for December and July 2005 to further examine reasons for discrepancies in CCSM-derived and observed soil-temperature climatologies. In this case study, we assume the soil temperatures simulated by WRF as a reference dataset from which we determine the "reference" regional averages. These reference averages are compared to the regional averages determined from WRF data at the sites of the network used in PaiMazumder et al. (2008) and to four randomly, but spatially even distributed artificial networks to assess exemplarily potential contribution of the historic network to the discrepancies found by these authors.

The high differences (1.8 K) between regional averages of soil temperature obtained from the historic network and "reference" in December suggest that the network design may affect gridded observational averages more in December than in July. This means

that the high discrepancies between CCSM-derived and observation-based gridded soil-temperature climatologies can be explained by the network design in winter.

PaiMazumder et al. (2008) also showed that in December, biases between simulated and observed soil temperature reach up to 6 K at 0.2 m depth; about 2.5 K bias may result from incorrect simulation of observed forcing. Our study shows that about 2 K bias may be explained by uncertainties due network density in winter. This means that about 1.5 K bias may result from measurement errors and/or model deficiencies.

In July, biases caused by the historic network are higher than biases in CCSM found by PaiMazumder et al. (2008). Hence, we have to conclude that in summer, CCSM performs well for simulating soil temperature. On the contrary, in winter, biases in CCSM can only partially be explained by uncertainties due to network density.

Similar results are found for RMSE and SDE in winter whereas in summer, RMSEs for the historic network are lower than RMSEs in CCSM by 1 K, on average. Hence, the discrepancies between CCSM-simulated and observation-based gridded soil-temperature climatologies in winter can be explained by incorrect simulation of atmospheric forcing as well as network design. Thus, improvement of soil-model physics is essential for better winter soil temperature simulation.

From this case study, it can also be concluded that the historic network always fails to capture the "reference" regional soil temperature averages with high biases, RMSEs and SDEs in both months. On the contrary, the randomly distributed 500-, 400- and 200-site networks capture the "reference" regional soil-temperature averages well at all layers. These networks also capture the diurnal variation of soil temperature well in the upper soil. Hence, our case study suggests that randomly distributed networks of 200-sites or more reliably reproduce acceptable regional averages of soil temperatures for Siberia. However, maintenance of such networks may be expansive because many of the sites would not be easily accessible in a remote area like Siberia. Future studies should examine the general robustness of the influence of the network density and design.

Acknowledgments

We thank U.S. Bhatt, M.E. Brown, G. Kramm, S.E. Porter, J.E. Walsh and the anonymous reviewers for fruitful comments, C. Swingly for help with the ERA40 data, B. Rudolf for access to the precipitation data, T. Zhang, A. Etringer, and J. McCreight for providing the gridded soil-temperature and snow-depth data and J. Miller for installing CCSM3 at ARSC. This research was supported by the International Arctic Research Center, UAF, under the auspices of the NSF cooperative agreement OPP-0327664and from EPSCoR grant 0701898. ARSC and NCAR provided computational support.

References

Anthes, R.A. 1983. Regional models of the atmosphere in middle latitudes. *Mon Wea Rev* 111: 1306-1335.

Anthes, R.A., Kuo, Y.H., Hsie, E.Y., Low-Nam, S. & Bettge, T.W. 1989. Estimation of skill and uncertainty in regional numerical models. *Quart J Roy Meteorol Soc* 111: 763-806.

Briegleb, B.P., Bitz, C.M., Hunke, E.C., Lipscomb, W.H., Holland, M.M., Schramm, J.L. & Moritz, R.E., 2004. Scientific description of the sea ice component in the Community Climate System Model, version three. NCAR Technical Note NCAR/TN463-STR, pp 70.

Collins, W.D., Bitz, C.M., Blackmon, M.L., Bonan, G.B., Bretherton, C.S., Carton, J.A., Chang, P., Doney, S.C., Hack, J.J., Henderson, T.B., Kiehl, J.T., Large, W.G., McKenna, D.S., Santer, B.D.& Smith, R.D. 2006a. The Community Climate System Model: CCSM3. *J Climate* 19: 2122-2143.

Collins, W.D., Rasch, P.J., Boville, B.A., Hack, J.J., McCaa J.R., Williamson, D.L., Briegleb, B.P., Bitz, C.M., Lin, S.J. & Zhang, M. 2006b. The formulation and atmospheric simulation of the Community Atmosphere Model: CAM3. *J Climate* 19: 2144-2161.

Dai, Y., Zeng, X., Dickinson, R.E., Baker, I., Bonan, G., Bosilovich, M., Denning, S., Dirmeyer, P., Houser, P., Niu, G., Oleson, K., Schlosser, A. & Yang, Z.L. 2003. The Common Land Model (CLM). *Bull Amer Meteor Soc* 84: 1013-1023.

Lawrence, D.M. & Slater, A.G. 2007. Incorporating organic soil into a global climate model. *Clim Dyn* doi: 10.1007/s00382-007-0278-1.

Li, Z., Bhatt, U.S., & Mölders, N. 2008. Impact of doubled CO_2 on the interaction between the regional and global water cycle in four study regions. *Clim Dyn* 30: 255-275.

Mitchell, T.D., Carter, T.R., Jones, P.D., Hulme M. & New, M. 2004. A comprehensive set of high-resolution grids of the monthly climate for Europe and the globe: the observed record (1901-2000) and 16 scenarios (2002-2100). Tyndall Centre for Climate Change Research, pp55.

Mlawer E.J., Taubman, S.J., Brown, P.D., Iacono, M.J. & Clough, S.A. 1997. Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J Geophys Res* 102D: 16663-16682.

Mölders, N. & Walsh, J.E. 2004. Atmospheric response to soil-frost and snow in Alaska in March. *Theor Appl Climatol* 77: 77-105.

Mölders, N. & Romanovsky, V.E. 2006.Long-term evaluation of the Hydro-Thermodynamic Soil-Vegetation Scheme's frozen ground/permafrost component using observations at Barrow, Alaska. *J Geophys Res* 111: D04105, doi: 10.1029/2005JD005957.

Narapusetty, B., & Mölders, N. 2005. Evaluation of snow depth and soil temperature predicted by the Hydro-Thermodynamic Soil Vegetation Scheme (HTSVS) coupled with the PennState/NCAR Mesoscale Meteorological Model (MM5). *J Appl Meteor*, 44: 1827-1843.

Nicolsky, D.J. & Romanovsky, V.E. 2007.Improved modeling of permafrost dynamics in a GCM land-surface scheme. *Geophys Res Lett* 34: L08501, doi:10.1029/2007GL029525.

Oleson, K.W., Dai, Y., Bonan, G.B., Bosilovich, M., Dickinson, R., Dirmeyer, P., Hoffman, F., Houser, P., Levis, S., Niu, G.Y., Thornton, P., Vertenstein, M., Yang, Z.L. & Zeng, X. 2004. Technical description of Community Land Model (CLM). Technical Report NCAR/TN-461+STR, National Center for Atmospheric Research, Boulder, CO. 80307-3000, pp 174.

PaiMazumder D., Miller, J., Li, Z., Walsh, J.E., Etringer, A., McCreight, J., Zhang, T. & Mölders, N. 2008. Evaluation of Community Climate System Model soil temperatures using observations from Russia. *Theor Appl Climatol* 94: 187-213.

Skamarock, W.C., Klemp, J.B., Dudhia, J., Gill, D.O., Baker, D.M., Wang, W. & Powers, J.G. 2005. A description of the advanced research WRF version 2. NCAR Technical Note, NCAR/TN-468+STR, pp 88.

Smirnova T.G., Brown, J.M. & Benjamin, S.G. 1997. Performance of different soil model configurations in simulating ground surface temperature and surface fluxes. *Mon Wea Rev* 125: 1870-1884.

Smirnova T.G., Brown, J.M., Benjamin, S.G. & Kim, D. 2000. Parameterization of cold season processes in the MAPS land-surface scheme. *J Geophys Res* 105D: 4077-4086.

Thompson G., Rasmussen, R.M. & Manning, K. 2004. Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part I: Description and sensitivity analysis. *Mon Wea Rev* 132: 519-542.

Zhang, T., Osterkamp, T.E. & Stamnes, K. 1996. Influence of the depth hoar layer of the seasonal snow cover on the ground thermal regime. *Water Resour Res* 32: 2075-2086.

Zhang, T., Barry, R.G. & Gilichinsky D. 2001. Russian historical soil temperature data. National Snow and Ice Data Center, Boulder, CO, http://nsidc.org/arcss078.html/.

Table 6.1 Monthly averages of bias and RMSE for the historic and 200-site-networks for upper and deeper soil

Networks		Historic network		200-site-network	
Month	Layer	bias	RMSE	bias	RMSE
July	Upper	0.6	0.7	0.02	0.1
	Deeper	1.7	1.9	0.3	0.3
December	Upper	1	1	0.1	0.2
	Deeper	1.2	1.2	0.2	0.3

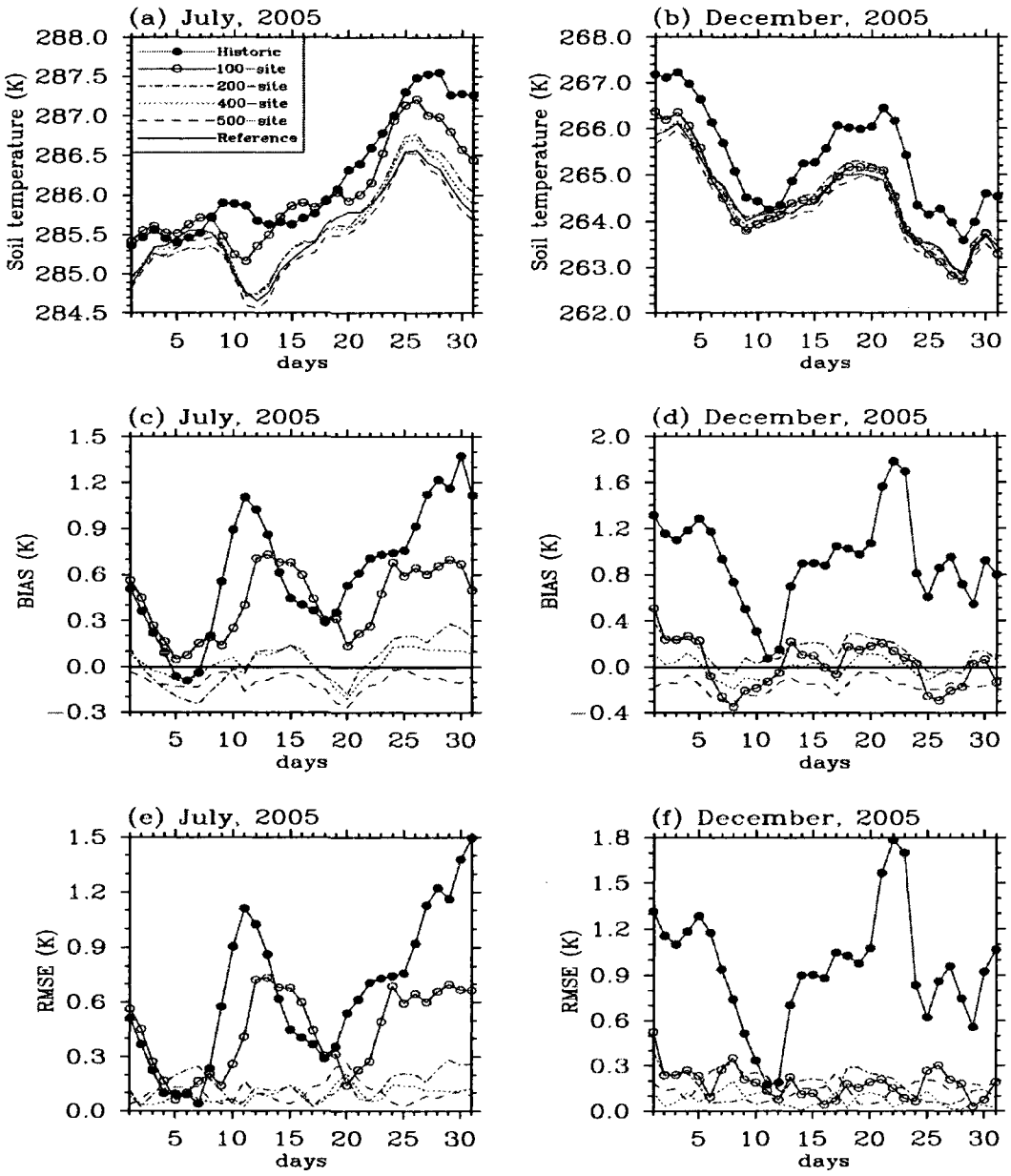


Fig. 6.1 Temporal behavior of (a) regionally averaged soil temperature at 0.2 m depth obtained from the "reference", 500, 400, 200, 100-site-network and the historic network for July, 2005 and (b) December, 2005. Biases for (c) July and (d) December and RMSE for (e) July and (f) December between the "reference" and 500, 400, 200, 100-site-network and the historic network. Note that in (c) to (f), labels on the y-axis differ. In (c) and (d), the thick line serves to better visualize the positive and negative bias.

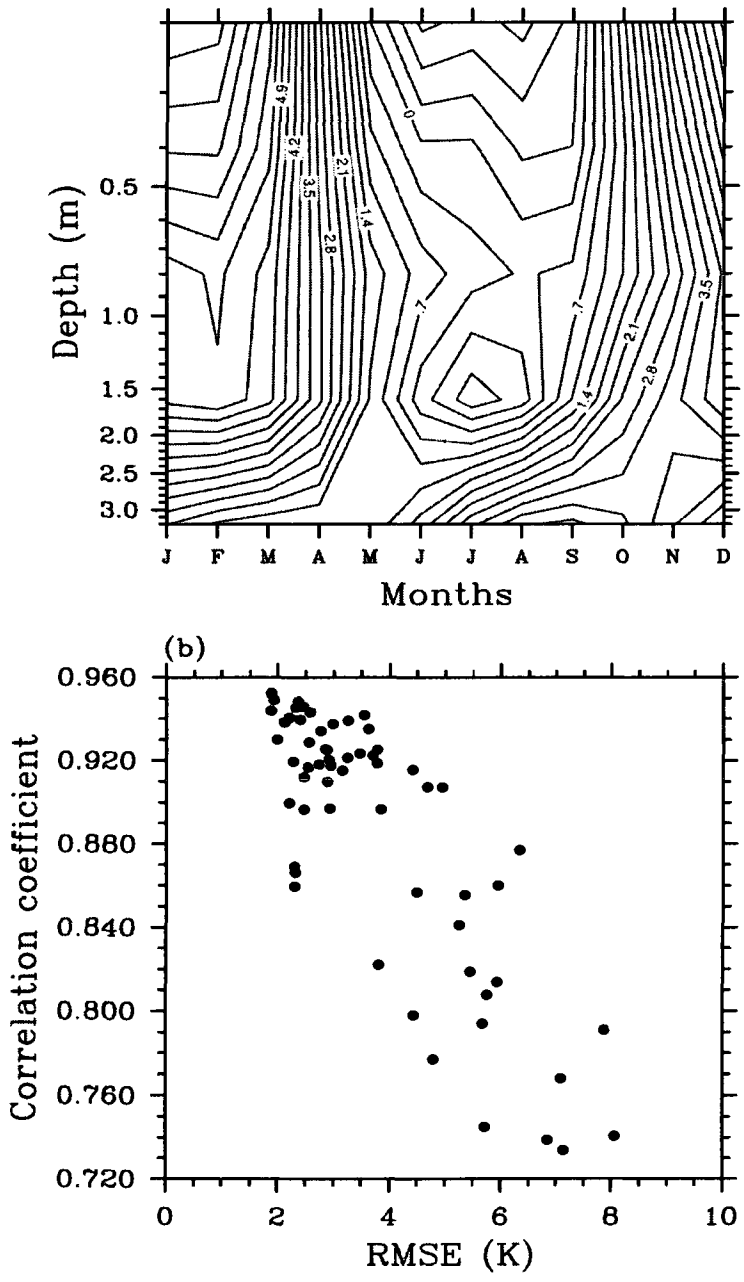


Fig. 6.2 (a) Contour plot of temporal behavior of biases with depths for the first climatology (1951-1980). (b) RMSE vs. correlation coefficient for the first climatology for all 12 months at 0.2 m, 0.4 m, 0.8 m, 1.6 m, and 3.2 m depth.

Chapter 7 Conclusions

Soil-temperatures simulated by the soil/permafrost model within the fully coupled Community Climate System Model version 3 are evaluated using 50 years (1950-2000) of Russian soil-temperature measurements that were gridded to the same resolution as the CCSM3 simulations. Evaluation of CCSM3-simulated near-surface air temperature, cloud, precipitation and snow depth with respect to their observations, results from sensitivity studies performed with the CCSM3 as well as simulations performed with the Weather Research and Forecasting model over Russia are used to assess (1) errors resulting from the soil/permafrost model itself, (2) errors from simulated atmospheric forcing due to running the soil/permafrost model in a fully-coupled mode with in the CCSM3 and (3) errors from inaccuracies in gridded soil-temperature observations that stem from density and/or design of the observational network.

The fully-coupled CCSM3 captures the phase of soil-temperatures well in the upper soil, and the annual average soil-temperatures reasonably well at all depths. However, it fails to capture the amplitude with higher (lower) soil-temperatures at all depths in winter (summer) than the observations with a better performance in summer than winter. In spring and fall, the simulated and observed soil-temperatures agree the best. The higher error in winter than in summer suggests that (1) the CCSM3 simulates warm soil processes of the active layer better than frozen soil processes and (2) the CCSM3's constant-heat-flux lower boundary condition represents actual conditions better in summer than winter, i.e. there seems to be a gradient in the winter flux. Therefore, future development should the address the lower boundary condition formulation.

The higher values of systematic errors in mountainous and lake-rich areas in the CCSM3-simulated soil-temperature are caused by consistent misinterpretation of physical, geometrical, and numerical aspects in CCSM3, and from the coarser observational network in these areas. Discrepancies between CCSM3 and real terrain elevation and the treatment of subgrid-scale heterogeneity, cloud and precipitation parameterizations, and pipe heat conduction cause these systematic errors. The first three

reasons clearly relate to the coupling. Systematic errors from pipe conduction affect both the offline and online evaluations. Other systematic errors stem from the fact that the network was in agricultural land, i.e. good soil is preferred. The higher values of random errors in mountainous and lake-rich areas in the CCSM3-simulated soil-temperature may result from higher observational errors due to blowing snow.

Since near-surface air temperature, cloud fraction, precipitation and snow depth can drive soil-temperature, the simulated values of these quantities are also evaluated with various observations. CCSM3 tends to overestimate winter and underestimate summer near-surface air temperature but overall CCSM3 simulates ERA40 near-surface air temperature reasonably well. Discrepancies in CCSM3-simulated near-surface air temperatures significantly correlate with the discrepancies in CCSM3-simulated soil-temperatures. Evaluation of precipitation by means of GPCC data shows that the CCSM3 fails to simulate summer precipitation adequately. Discrepancies in precipitation also significantly correlate with soil-temperature discrepancies in summer. Evaluation of snow depth reveals that the incorrect simulation of the fall snowfall amount and the snow depth explains some of the errors in the simulated soil-temperature in winter. Cloud fraction discrepancies also correlate significantly with soil, near-surface air temperature, and precipitation discrepancies; the correlation is especially high in summer. Based on these correlations we conclude that convective clouds and precipitation parameterization shortcomings may be the main reason for the underestimation of summer soil-temperatures. From all these findings it can be concluded that incorrectly-simulated atmospheric forcing contributes to, but is not the main reason for discrepancies between CCSM3-simulated and gridded observed soil-temperature. Thus, improving clouds, precipitation parameterizations and surface heterogeneity parameterizations and using a finer model resolution that resolves terrain height and coastlines better than T42, are urgent need for improving the soil-temperature simulations by the fully coupled CCSM3.

The sensitivity studies with altered sand and clay indicate that soil characteristics contribute notably or even significantly to the errors in the simulated soil-temperature climatologies. The fact that decreasing or increasing sand percentage yields no clear

result further suggests that better or worse results may occur by accident. Systematic error may result from using mineral soil instead of organic soil physical parameters. Large areas of Russia, however, have organic soils at least in the upper soil layers while CCSM3 assumes mineral soils. Due to noticeable differences in the thermal and hydraulic properties of organic and mineral soil, the inclusion of organic material is an urgent need for the soil/permafrost model. A sensitivity study with altered albedo suggested that uncertainty in surface albedo may have an impact on simulated soil-temperature climatologies and hence on the permafrost distribution and the active layer depth in the fully-coupled model. A sensitivity study assuming slightly different landuse reveals that difference in landuse may also have an impact on simulated soil-temperature climatologies. Based on these sensitivity studies, one has to conclude that some discrepancies between CCSM3-simulated and gridded soil-temperatures are due to differences in soil and vegetation types assumed and actual soil and vegetation characteristics. Thus, one must conclude that better data on soil and vegetation types and functions on the temporal change of vegetation parameters will provide potential for better soil-temperature simulations. Inclusion of organic soils in permafrost simulation is an urgent need to be addressed in future research.

It is well known that gridded observational data typically has some uncertainty due to interpolation procedure. Most of the Russian soil-temperature sites are associated with long-term agricultural monitoring stations and are designed with accessibility and ease of maintenance in mind. Consequently, the observational network is not uniformly distributed and biased with respect to terrain elevation, soil type (fertile soils) and land-cover type (agriculturally used land). To examine whether the density and/or design of networks may introduce some additional uncertainty in the gridded datasets of regional averages based on station data, the WRF model is used to provide a dataset for soil-temperatures that is considered as a "reference" for determination of regional averages. Data from the real, also called historic, network used in the CCSM3 evaluation and randomly, but spatially even, distributed artificial networks are compared to the "reference" gridded data. Networks with 200 or more randomly distributed sites reliably

reproduce the reference regional averages of soil-temperature while the real network has difficulties in capturing the reference regional averages. This means that non-random network design introduces substantial uncertainty in gridded data and that networks with randomly distributed sites might only need about half of the points of the non-random distributed real or historic network over Russia to determine gridded data. However, the maintenance of such networks with randomly distributed sites can be extremely expensive, especially when sampling is to be performed over several decades, because many of the sites would be difficult to access in remote areas. Based on these findings, one has to conclude that evaluation studies using long-term data taken for other purposes require intelligent concepts to derive reasonable conclusions and to avoid tuning models towards biased network data. The results also suggest that downscaling/interpolation of observations may be problematic in complex terrain. An important finding is also that if a less number of randomly placed sites provide better results than a large number of ill placed sites, intelligent network design may save costs and increases the knowledge.

The biases between CCSM3-simulated and observed soil-temperature are as high as 6 K at 0.2 m depth in winter. About 2.5 K of the bias may result from the incorrect simulation of the observed forcing and about 2 K of the bias may be explained by uncertainties due network density (Fig. 7.1). This means that about 1.5 K winter-bias may result from measurement errors and/or model deficiencies. In July, biases caused by uncertainties due network density are higher than biases in the CCSM3 soil-temperature simulation (Fig. 7.1). Hence, it can be concluded that in summer, CCSM3 performs well for simulating soil-temperature. On the contrary, in winter the high discrepancies between the CCSM3-derived and observation-based gridded soil-temperature climatologies can be explained by the observational network design. The results from this analysis indicate that the scientific community has to work on obtaining better, more representative observational data.

From this study, it can be concluded that the performance of a soil/permafrost model fully coupled with a climate model depends partly on the soil/permafrost model itself, the accuracy of the forcing data provided by the climate model and the design of the

observational network. Improving parameterizations and model resolution are urgent needs for improving soil-temperature simulations and, hence, any assessments of permafrost thawing by the fully coupled CCSM3. Non-randomly distributed networks may provide wrong impressions on regional climate and its changes. Therefore, future networks should be designed in a more spatially random method. Until long term data from randomly designed networks become available, one will have to use whatever data are available. Furthermore, data from non-randomly distributed sites are only available for the past. Therefore, when evaluating climate models using gridded data from non-randomly distributed networks, one has to develop intelligent strategies to guarantee meaningful conclusions on model performance and for model improvement.

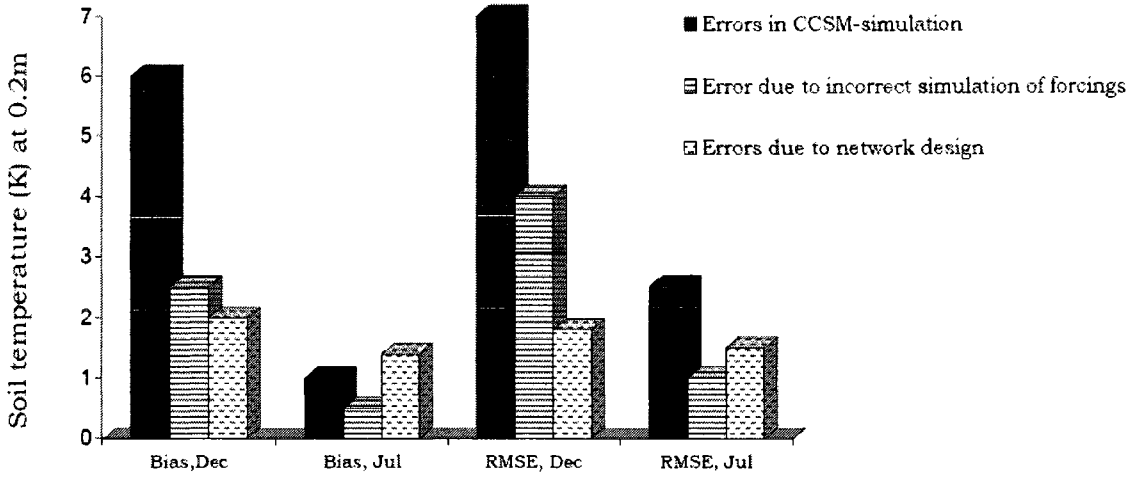


Fig. 7.1 Biases and RMSEs between CCSM3-simulated and observed soil-temperature, amount of biases and RMSEs explained by incorrect simulation of observed forcing by CCSM3 and uncertainty due to network design for July and December.

Appendix A Contributions of thesis chapters

A.1 Chapter 4

The key topic of this chapter was adapted from Professor Mölders' NSF proposal under grant OPP-0327664. Gridded data of soil-temperature and snow-depth were provided by Zhang T, Etringer A and McCreight J (National Snow and Ice Data Center). Jack Miller installed CCSM3 on the ARSC supercomputer. The literature research, simulations and figures were prepared by Debasish PaiMazumder. The reference simulation was performed by Professor Mölders. Dr. Li performed the simulation wherein forest is randomly replaced by grass and/or crops in some Russian grid-cells (total area change <1%). Professor Mölders also helped in the physical interpretations and refining of the text and the figures. Professor Walsh also provided fruitful comments during the editing process.

A.2 Chapter 5

The experimental design was provided by Professor Mölders. The literature research, text, simulations and figures were prepared by Debasish PaiMazumder. Professor Mölders helped PaiMazumder in the physical interpretations and refining of the text and the figures.

A.3 Chapter 6

The key topic of this chapter was adapted from Professor Mölders' NSF proposal under grant OPP-0327664, and ARC0652838, and EPSCoR-grant 0701898. The literature research, text, simulations and figures were prepared by Debasish PaiMazumder. Professor Mölders helped PaiMazumder in the physical interpretations and refining of the text and the figures.